

THIS WEEK

EDITORIALS

ANTHROPOCENE Time to call time on human impact on planet **p.254**

WORLD VIEW Protect and pay for seagrass, marsh and mangrove **p.255**



POSER Cuttlefish finds imitation the sincerest form of protection **p.256**

There's a time to be critical

An accusation that referees are too demanding and editors too supine demands a response. Authors, editors and referees all have lessons to learn.

Last week one of our editors received the following from a referee of a paper currently under assessment:

"I guess the issue with this kind of paper is that there are an almost limitless number of changes/additions that could be made, especially considering the complexity of the data presented here. I suspect that this paper might run into a few reviewer 'issues' as it covers so much ground. In my review I have tried to be cognisant of your 27 April *Nature* article ('End the wasteful tyranny of reviewer experiments') and as such give this a 'yes' vote pending revisions."

In the same week, we received a note from another reviewer to the effect that the "tyranny of reviewer experiments" had significantly increased the impact of the claims made in a manuscript he assessed, and he hoped that the authors would agree that the further work was worth the effort.

Clearly, some targets of the *Nature* article have taken note of it. In brief, that column, by Hidde Ploegh at the Massachusetts Institute of Technology in Cambridge, argued that referees too often ask for more experiments, and that editors too passively tend to pursue such requests (see *Nature* **472**, 391; 2011).

But for the paper mentioned above, the question of whether further work is required is still open until the editor decides otherwise. Our editors must ask themselves: would further work lift the paper over a threshold of robustness or significance that justifies publication in *Nature*, or is it already sufficient? And have other referees differing views about this?

In resolving these questions, the editor will discuss the paper with colleagues and also with the referees.

The accusation that editors are too passive was not specifically directed at *Nature*, but we take it seriously. We could too easily discount it on several grounds. Surveys of our published authors, as well as general surveys of scientists conducted independently, overwhelmingly support the view that papers have gained in their passage through peer review. Critics do not realize how much discussion and critical assessment underpins our editorial decisions. And without question, the ever-increasing pressure to publish is far too often leading authors to submit papers that would gain substantially in scientific significance with some further work.

It is important also to acknowledge that our referees generally put in very substantial amounts of labour on behalf of their fellow scientists, and make constructive suggestions that ensure that some of the extraordinary claims that *Nature* publishes are backed by the necessary evidence.

Nevertheless, a more reflective response is also required.

At *Nature* and at the *Nature* research journals, our teams of staff editors are expected to make their own conclusive judgements about a paper's position below or above their journal's threshold, and will often overrule referees' expectations in this respect in either direction. For example, we may decide that even if a paper lacks a new

insight into mechanism, it represents a sufficient resource in the novelty of its data or technique to make a significant impact on the discipline. Conversely, we may decide that an additional piece of work would greatly increase a paper's range or depth of impact, and make that a condition of publication — we hope to the ultimate benefit of the community and the authors themselves (see *Nature* **463**, 850; 2010).

"Referees generally put in very substantial amounts of labour on behalf of their fellow scientists."

But our editors do not necessarily have the expertise to judge whether, for example, an application of a novel technique or reagent has been adequately validated. Authors are free to challenge a request for more work in these circumstances, and an editor may seek technical advice from another expert to resolve the matter.

Spurred by this discussion, we looked back at recent decisions. We soon found several cases in which, with technical guidance where necessary, we overruled a referee's request for additional work — for example, when the editor felt that, contrary to a referee's assertion, the gain in robustness would not be sufficient to justify the effort and delay.

What lessons can be learnt, therefore? By authors: in the interests of robustness and genuine impact, resist the pressure to publish prematurely. By editors everywhere: don't be supine in the face of referees' requests.

And above all, by referees: please don't ignore any impulse to demand more, but be self-critical too. ■

Getting personal

Targeted therapies work, but need help to fulfil their potential.

Biology is like economics, participants at a European Commission meeting on personalized medicine in Brussels heard last week: they are both complex and neither is properly understood. The view struck a chord with attending scientists and health-care economists, who felt that personalized medicine should be happening, and didn't understand why, mostly, it isn't.

Personalized medicine aims to use the latest genomic knowledge and technologies to tailor treatments to individuals. Pivotal to the field are drugs that have been designed to hit a particular molecular pathway that has gone wrong in a disease. The European Medicines Agency has already approved around 15 such drugs for cancer therapy

and is set to approve several more in the next year or so.

The personalized approach faces two major problems: complex biology and complex economics. The pathway involved is often not well understood, and most targeted drugs are so expensive that health-care systems and insurance companies don't want to pay for them, even if they reduce waste and should therefore save on overall treatment costs in the long term. The drug gefitinib, for example, costs around €20,000 (US\$28,000) per patient and targets the *EGFR* pathway, which is disrupted in fewer than 15% of patients with lung cancer. What's more, targeted drugs need to be accompanied by diagnostic tests to identify suitable patients, yet many health-care systems have no mechanism to pay for the tests. The result is an absurd situation in which expensive drugs can be prescribed without testing, and therefore to some patients who will gain no benefit.

As arguments about the value of personalized medicine rage around the world, France has found its own solution — at least for cancer, where molecular medicine is most advanced. In 2005, the country said it would pay for the treatment of every citizen shown to be likely to benefit from targeted drugs. Its National Cancer Institute set up 28 platforms for molecular genetics at university hospitals and cancer centres with expertise in both molecular and pathological analysis. Biopsies of cancerous tissue from patients all over France are sent to these platforms for a battery of 20 or so genetic tests. If the tissue displays a genetic signature in any molecular pathway targeted by one of the drugs, the patient gets treated with it. The platforms develop the tests themselves, and are already working on a test to accompany a drug that researchers hope will be approved this year for melanoma. Targeted drugs now account for 57% of France's cancer-treatment budget. The Czech Republic has a similar system.

The model seems to work. The French platforms have so far tested samples from around 15,000 people with lung cancer for alterations in the *EGFR* pathway. Just over 1,700 patients tested positive and were given gefitinib until they stopped responding (an average of

38 weeks). That has cost France €35 million. Had all 15,000 patients been given an eight-week course of gefitinib just to see whether they would respond, it would have cost the nation another €69 million — with no extra benefit.

Some assessments, however, have concluded that personalized drugs do not offer enough benefit to justify the cost. It will not be easy to persuade the spectrum of state health systems and health-insurance companies that personalized medicine makes economic sense. Understandably, they will want a lot more evidence that it works.

Much reluctance also seems to come from a medical profession

“As arguments about the value of personalized medicine rage around the world, France has found its own solution.”

unused to needing genetic tests to select patients and from inflexible bureaucratic systems. The European Commission's health directorate could help by encouraging European countries to harmonize their health-technology assessments, or even by issuing its own (non-binding) conclusions on which targeted drugs it considers cost-effective. And the commission's research directorate could provide greater support for efforts to translate the results of pre-

clinical research on molecular pathways into the clinic, which it plans to do in its 2012 call for proposals.

Amid the excitement and attention paid to cancer, it is crucial to remember that other conditions — such as psychiatric disorders — carry just as great a societal burden, yet remain too poorly understood to benefit. The research directorate has enabled a great deal of fundamental research on animal models designed to understand such complex conditions, and it must continue to do so, in parallel with its translational efforts. We are at the beginning of personalized medicine in the clinic. But we are also just starting to understand the mechanisms behind most of the diseases that are likely to gain the most. ■

The human epoch

Official recognition for the Anthropocene would focus minds on the challenges to come.

Geologists are used to dealing with heavy subjects, so who better to decide on one of the more profound debates of the time: does human impact on the planet deserve to be officially recognized? Are we living in a new geological epoch — the Anthropocene?

This is no idle conundrum. Although the term has long been used informally to refer to the current, human-dominated phase of Earth's history, a working group of the International Commission on Stratigraphy, the body that defines the divisions of geological time, is studying the case for making it official (see *Nature* 473, 133; 2011).

The Anthropocene would be a peculiar addition to the geological timescale. So far, it is more a prediction than a fact of Earth's history, because many of its defining features are only starting to register in the rock record. And the driving force behind the geological transition it labels is not a continental rearrangement, massive volcanism or an extraterrestrial impact — forces that have reshaped the planet in the past. Yet the Anthropocene does deserve proper recognition. It reflects a grim reality on the ground, and it provides a powerful framework for considering global change and how to manage it.

Human activity is set to leave an indelible mark on the geological record. Deforestation, mining and road building have unleashed tides

of sediment down rivers and onto the ocean floor. Fossil-fuel use and land clearance have already emitted perhaps a quarter as much carbon into the atmosphere as was released during one of the greatest planetary crises of the past, the Palaeocene–Eocene Thermal Maximum 55 million years ago. Now, as then, corals and other organisms are recording a global carbon-isotope shift. The increasing acidification of the oceans as they absorb carbon dioxide will dissolve carbonate from deep sediments, and what is likely to be the sixth great mass extinction in Earth's history will gather speed, adding vivid new markers to the record.

But is it too soon to declare an end to the Holocene, the stable, largely benign epoch that has lasted just 11,700 years — a heartbeat in geological time? What impact will an official change in the geological timescale have on the funding and status of Holocene studies? And is it wise for stratigraphers to endorse a term that comes gift-wrapped as a weapon for those on both sides of the political battle over the fate of the planet?

The scale of the changes already under way and the real value of a unified approach to studying human influences on the planet should surely quash these concerns. The Anthropocene is defined not just by climate change or extinctions, but by a linked set of effects on Earth and its biosphere, from perturbations in the nitrogen cycle to the dispersal of species around the globe. Official recognition of the concept would invite cross-disciplinary science. And it would encourage a mindset that will be important not only to fully understand the transformation now occurring but to take action to control it.

Humans may yet ensure that these early years of the Anthropocene are a geological glitch and not just a prelude to a far more severe disruption. But the first step is to recognize, as the term Anthropocene invites us to do, that we are in the driver's seat. ■

► NATURE.COM
To comment online,
click on Editorials at:
go.nature.com/xhunqv



Add coastal vegetation to the climate critical list

Forests are protected, but carbon sinks in mangroves, seagrasses and marshes are ignored. Margareth da Silva Copertino wants Brazil to change that.

My country, Brazil, is home to 80% of the remaining Amazon rainforest, and has rightly worked to find ways to sustainably manage and commercialize these stocks of forest carbon. However, like most countries with long coastlines, Brazil has so far missed the opportunity to value and protect another important carbon store: its mangroves, seagrasses and tidal marshes.

The 9,000-kilometre vibrant and productive Brazilian coastline is covered with vegetated ecosystems that together contain hundreds of millions of tonnes of such carbon, at least. Brazil is home to the third-largest mangrove area in the world and has more than 20,000 hectares of seagrasses near tropical reefs and in coastal lagoons.

Why aren't these systems recognized as vital pieces of the climate-change puzzle? They cover just 0.5% of marine areas, but are among the largest carbon sinks in the ocean. Typically, they store up to 15 times more carbon per hectare than terrestrial soils, absorbed over hundreds or even thousands of years. And these coastal systems sequester carbon 10–50 times faster than terrestrial forests.

As the United Nations REDD+ scheme to protect the 'green' carbon stocks in tropical forests develops, it is time to broaden the reach of such mechanisms so that they also value and protect coastal stocks of 'blue' carbon.

In February, I took part in a scientific workshop in Paris to evaluate such mechanisms and offer policy-makers the information and advice they need to make them happen. The event was a meeting of the International Working Group on Coastal Blue Carbon, formed jointly by Conservation International, The International Union for Conservation of Nature and the Intergovernmental Oceanographic Commission of the UN Educational, Scientific and Cultural Organization. It concluded that coastal carbon deposits should be taken into account in national emission inventories and in the processes and mechanisms of the UN climate framework. The group will continue to meet over the next two years and will urge policy-makers to recognize the importance of blue carbon.

They must do this soon, because coastal areas are among the most threatened ecosystems on Earth. Between 30% and 50% of mangroves have disappeared in the past 50 years; about 30% of the world's seagrasses are gone; and half of the global coverage of salt marshes has been destroyed. That loss is continuing and in many places accelerating — 2% of those important coastal systems are lost each year. That is four times the estimated rate of tropical-forest loss.

Degradation and destruction of blue-carbon systems results in immediate and continued emissions. Studies by Duke University in

Durham, North Carolina, estimate that emissions from such clearing result in up to 900 million tonnes of carbon dioxide emissions per year, roughly equal to the annual CO₂ emissions from energy consumption and industry for the whole of Germany. That is about 10–20% of the emissions from deforestation globally, or 2% of all anthropogenic greenhouse-gas emissions.

Brazil would be a good place to test new mechanisms to value and conserve blue carbon. The country has about 200 protected areas along the coastline, spanning different latitudes and ecosystems. But these cover just 20% of the country's total coastal territory, and represent only one-quarter of the area highlighted by the Brazilian government as a conservation priority. The country's mangroves,

salt marshes and seagrasses are under mounting pressure from a combination of intense human activities, increasing coastal development (about one-fifth of the Brazilian population lives by the coast), agricultural run-off, pollution and intensive aquaculture. These rising threats require the urgent application of mechanisms to increase the monetary value of coastal habitats, promote their conservation and avoid further degradation.

There is more to protecting coastal ecosystems than global recognition of their importance, however. We need focused and coordinated research, as well as scientific data collection, to build financial mechanisms that value these ecosystems as tools to reduce greenhouse-gas emissions.

But local policies and regulation are vital too. Developing nations must support small landowners and the livelihoods of local com-

munities, expand the law-abiding 'responsible' fraction of economic sectors, improve law enforcement, effectively manage protected areas and recover the many degraded ones. In Brazil, as in other countries, we need to stop the destruction of existing mangroves, tidal marshes and seagrasses.

Effectively accounting for the carbon in coastal systems has the potential to transform the management and conservation of coastal areas on both the global and local scales. For the sake of Brazil and the world, my country should push for the role of oceans and their coastal ecosystems to be included in UN climate talks this year, alongside rainforests. Brazil has already played an important part in many of these discussions, but it could be a true leader if it were to recognize the significant potential hidden in its long, blue coast. ■

Margareth da Silva Copertino is a lecturer in biological oceanography at the Institute of Oceanography, Federal University of Rio Grande (FURG), Rio Grande, Brazil.
e-mail: doccoper@furg.br

DEGRADATION AND
DESTRUCTION OF
BLUE-CARBON
SYSTEMS
RESULTS IN
IMMEDIATE
EMISSIONS.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/a5rgal

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

MICROBIOLOGY

Starved cells turn on themselves

Bacteria are so essential to mammalian digestion that without them, some gut cells break down their own components to obtain energy.

Scott Bultman at the University of North Carolina at Chapel Hill and his colleagues studied the impact of gut bacteria on the metabolism of mice. They found that in mice lacking any bacteria, colon cells are energy deprived and undergo autophagy, or 'self-eating'. Putting bacteria that produce butyrate — colon cells' main energy source — into the guts of these mice returned the cells' metabolism to normal.

A decrease in butyrate-producing bacteria in the gut, perhaps caused by dietary changes, could compromise colonic function, the authors say. This might contribute to higher rates of inflammatory bowel disease and colon cancer. *Cell Metab.* 13, 517–526 (2011)

ANIMAL BEHAVIOUR

Strike a pose and hide

Cuttlefish evade predators by matching not only their colours and patterns to the background, but also their postures.

Roger Hanlon and his team at the Marine Biological Laboratory in Woods Hole, Massachusetts, presented the common European cuttlefish

(*Sepia officinalis*; pictured left) with separate backgrounds containing stripes at different angles: horizontal, vertical, and diagonal. The animals raised their limbs to match the angles of the stripes, but didn't respond to a blank background.

When placed next to artificial algae in the lab, the animal struck a pose to mimic its neighbour (right). Similar behaviour was observed in natural habitats. The authors suggest that visual cues are

important for such creatures to adopt cryptic body postures and achieve maximum stealth. *Proc. R. Soc. B* doi:10.1098/rspb.2011.0196 (2011)

DRUG DESIGN

Designer proteins target flu

Proteins that bind to the 1918 pandemic influenza virus have been designed using computer modelling.

The viral surface protein haemagglutinin is essential to the flu virus's infection of human cells, making it an attractive drug target. David

Baker at the University of Washington in Seattle and his colleagues computed 'hot spots' — protein residues that can interact with haemagglutinin — from the 1918 virus on the basis of properties such as the predicted strength of their interaction with this protein. They then used computer algorithms to search a set of 865 protein structures for ones that could incorporate these hot spots, and came up with 88 proteins able to acquire at least two of them. When expressed in yeast, two of these proteins bound to haemagglutinin. After additional optimization, the researchers solved the



ZOOLOGY

Warblers of the underwater world

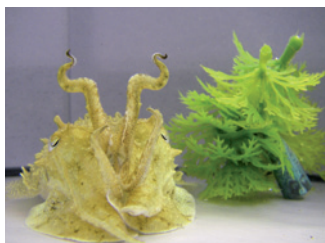
Many birds, mammals and amphibians vary the frequency and intensity of their vocalizations to expand their vocabulary. Aaron Rice, Bruce Land and Andrew Bass at Cornell University in Ithaca, New York, show that fish also use forms of 'acoustic nonlinearity', such as frequency jumps and biphonation — the simultaneous expression of two independent frequencies.

The authors recorded and analysed the vocal calls of three-spined toadfish (*Batrachomoeus*

trispinosus; pictured), which produce 'hoots' and 'grunts' by vibrating their swim bladders. Around 35% of the fish's calls had at least one form of nonlinearity. Severing the animals' vocal motor nerve stopped them producing these effects.

The fact that fish make complex vocalizations previously found only in four-limbed vertebrates suggests that there is a major selection pressure to produce innovation in acoustic signals.

Proc. R. Soc. B doi:10.1098/rspb.2011.0656 (2011)



molecular structure of one of these tailored proteins while it was bound to the flu haemagglutinin, revealing atomic-level accuracy in the designed interaction.

Science 332, 816–821 (2011)

IMMUNOLOGY

Blocking brain inflammation

Harmful brain inflammation triggered by a subset of immune cells can be quelled by the action of a hormone on an oestrogen receptor.

Microglia are immune cells that trigger inflammation in the central nervous system and carry an oestrogen receptor called ER β . Christopher Glass and Kaoru Saijo at the University of California, San Diego, and their team screened a panel of molecules that bind to ER β for their ability to block inflammation in microglia. They found that a few synthetic chemicals, as well as a natural steroid hormone called ADIOL, activate ER β , kicking off a cascade of reactions that ultimately prevents inflammation. ADIOL also protects mice from an autoimmune condition similar to multiple sclerosis.

The authors suggest that drugs that stimulate this pathway could be used to treat neurodegenerative and autoimmune diseases.

Cell 145, 584–595 (2011)

SEISMOLOGY

Awakening a fault line within

A large fault section off Sumatra that had been seismically dormant for more than 30 years has recently reawakened, thanks to a series of large earthquakes in the area during the past decade.

Kelly Wiseman at the University of California, Berkeley, and her team linked data derived from the Global Positioning System on surface motions from Sumatran stations with the known

geometry and mechanisms of recent quakes. They found that a 900-kilometre-long ‘backthrust’ — arising from the longer Sunda megathrust fault that caused the 2004 Indian Ocean earthquake and tsunami — produced a moderate quake in 2005 and another in 2009.

If a rare faulting event were to rupture the newly active thrust, it could produce a quake on the order of magnitude 8.5, and, potentially, a large tsunami, the authors suggest. *Geophys. Res. Lett.* doi:10.1029/2011GL047226 (2011)

PHYSICS

Achieving spin control

By combining an optical microscope with an atomic force microscope (AFM), researchers have imaged individual electronic spins with high resolution.

Amir Yacoby and his team at Harvard University in Cambridge, Massachusetts, implanted clusters of nitrogen ions into a diamond sample, creating individual spins. They applied a magnetic field gradient to the spins by passing the magnetized tip of the AFM over the sample. This allowed them to draw a map of the sample's individual spins in three dimensions. The authors showed that it would be possible to resolve spins just 9 nanometres apart.

The system could be used for studies of fundamental physics, because the set-up allows quantum control and manipulation of individual spins.

Nature Phys. doi: 10.1038/nphys1999 (2011)

AUTOIMMUNITY

Taming psoriasis with vitamin D

Vitamin D may ameliorate the symptoms of the inflammatory skin disease psoriasis by enhancing the production of a molecule that blocks the assembly of inflammatory complexes in the skin.

COMMUNITY CHOICE

The most viewed papers in science

NEUROSCIENCE

A search for depression genes

HIGHLY READ
on cell.com
10 April–10 May

A genome-wide analysis of more than 15,000 people has revealed an association between a gene and major depression.

Martin Kohli and Elisabeth Binder at the Max Planck Institute of Psychiatry in Munich, Germany, and their colleagues first compared the genomes of 353 patients with depression with those of 366 controls. They teased out a gene, *SLC6A15*, that was strongly associated with depression, and went on to replicate this finding in six other independent groups of patients. The gene encodes a transporter protein that moves certain amino acids across the cell membrane of neurons and may be involved in regulating the transmission of glutamate, a neurotransmitter.

The gene variant linked with depression was associated with reduced *SLC6A15* expression in the human hippocampus, as well as decreased volume of this brain region.

Neuron 70, 252–265 (2011)

Jürgen Schaubert and Robert Besch at the Ludwig Maximilian University in Munich, Germany, and their co-workers found that, under certain conditions, DNA in the cytosol of cultured skin cells activates immune complexes called inflammasomes that contain the protein AIM2. Elevated levels of this DNA and AIM2 expression were also found in skin cells from people with psoriasis. When normal skin cells were treated with the antimicrobial peptide cathelicidin LL-37, whose production in the skin is controlled by vitamin D, the peptide bound to cytosolic DNA, inhibiting the formation of AIM2-containing inflammasomes.

Stimulating cathelicidin production may be a promising approach for treating psoriasis, the authors suggest.

Sci. Trans. Med. 3, 82ra38 (2011)

MATERIALS SCIENCE

Diamond lighter than a feather

Aerogels are extremely porous and lightweight materials with a large surface area and many potential applications.

Peter Pauzauskie, now at the University of Washington in Seattle, and his colleagues have created a diamond version of the material (pictured), by squeezing an aerogel of amorphous carbon until it took on a crystalline structure.

The authors used high-pressure neon gas to fill and support the delicate carbon structure. They then zapped it with a laser that compressed and heated the gel, probably to more than 1,600 kelvin, until it became diamond.

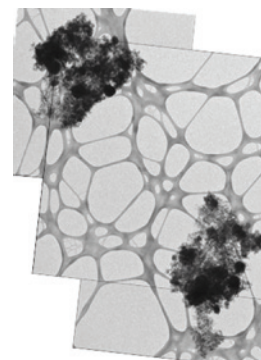
Diamond aerogels could be useful as antireflective coatings, thermal conductors and other materials, the authors say.

Proc. Natl Acad. Sci. USA doi:10.1073/pnas.1010600108 (2011)

NATURE.COM

For the latest research published by Nature visit:

www.nature.com/latestresearch



P. PAUZAUSKIE

SEVEN DAYS

The news in brief

POLICY

Nuclear plans in flux

Uncertainty over the future of nuclear power grew in several countries last week. Japan's prime minister Naoto Kan said his country would review its nuclear policy (for more, see page 263). In Germany, a leaked draft report from an ethics commission on safe energy, set up by Chancellor Angela Merkel, recommended shutting all nuclear plants by 2021. In the United States, inspections of nuclear plants by the Nuclear Regulatory Commission threw up a number of flaws. And European nuclear-safety regulators met in Brussels to announce details of stress testing for nuclear plants, but could not agree on criteria.

Fisheries reform

Drafts of the European Commission's proposed overhaul of Europe's fishing industry were leaked to the media last week. The commission wants to cut catch quotas so that by 2015, stocks are fished at the maximum yield that is sustainable. It also hopes to let fishermen buy and sell catch quotas, and to ban the practice of throwing back some caught fish. Some scientists worry about the expense of the research base and enforcement required to sustain and police such a system. The final proposal will be presented on 13 July. See go.nature.com/edjtqe for more.

US retains brains

In a bid to keep talented overseas researchers, the US Department of Homeland Security is giving more foreign science students extra time to search for work after graduation. Graduates can usually apply to stay in the United States for a year, receiving related postgraduate



INTERNATIONAL WOW COMPANY

Shale-gas fracking faces French ban

When the Oscar-nominated documentary *Gasland* showed people setting their tap water on fire (pictured), it thrust hydraulic fracturing into the spotlight. The technique, also known as 'fracking', in which high-pressure fluids are pumped into shale formations to fracture the rock and force out natural gas, has been accused of releasing methane into well water (hence, perhaps, the flammable tap water) and of polluting groundwater with toxic chemicals.

Now France could become the first country to ban the practice. On 11 May, the French parliament's lower house voted for a ban; the upper house will vote next month. Several areas in the United States have recently issued moratoriums on fracking, and a panel set up on 5 May by Department of Energy head Steven Chu will report in three months on how to improve its safety. See go.nature.com/wkizub for more.

training, before having to switch visas or leave the country. But on 12 May, the department added a range of science-related subjects — including neuroscience, drug design and environmental science — to a list of degrees eligible for a 29-month postgraduate stay.

Openness inquiry

The Royal Society in London is asking whether scientific data should be shared more openly with the public. On 13 May, it launched an inquiry into how scientific information should be managed "to improve the quality of research and build public trust". A working group chaired by Geoffrey Boulton, a glaciologist at the University

of Edinburgh, UK — and including the editor-in-chief of *Nature* — will release its conclusions in early 2012. It is accepting submissions of evidence until 5 August. See go.nature.com/w2di6n for more.

IPCC overhaul

The Intergovernmental Panel on Climate Change last week agreed to change its workings and governance. The reforms were decided at a general assembly in Abu Dhabi. See page 261 for more.

Arctic land rush

Cooperation, not conflict, was the message relayed by politicians at the Arctic Council's biennial meeting

in Nuuk, Greenland, on 12 May. The rush to claim seabed territory and oil and gas resources in the melting Arctic threatens to spark international disputes. But US secretary of state Hillary Clinton described the council as the 'preeminent intergovernmental body' for solving problems. The meeting saw the signing of the council's first legally binding treaty between Arctic nations — although it relates only to cooperation on search and rescue missions in the region.

Korean science belt

South Korea's 5.2-trillion-won (US\$4.8-billion) 'science belt' project will be based — unsurprisingly — in a region

already home to the country's leading science university, the Korean Advanced Institute of Science and Technology, and numerous industrial laboratories. Korean media said a site-selection committee had chosen the Daedeok research district, in the city of Daejeon, to host a planned 410-billion-won rare-isotope accelerator and 25 of the 50 laboratories of a new basic-science institute. The infrastructure is planned to be built by 2017.

FUNDING

Low odds at the NIH

Fewer than one in five research grant applications to the US National Institutes of Health (NIH) will gain funding in the 2011 fiscal year, according to Francis Collins, the agency's director. Predicted success rates of 17–18% would be “the lowest in history”, Collins told a Senate committee on 11 May. In 2010, the NIH's grant-application success rate was 20%. See go.nature.com/g9nh4t for more.

UK facilities cuts

The UK government has revealed cuts to planned scientific facilities, part of an effort to shrink the national deficit. Three projects have definitely been axed: a national supercomputing service called ARCHER; a planned upgrade

to the Rothera Research Station in Antarctica; and a £50-million (US\$80-million) centre for computational science at the Daresbury Science and Innovation Campus in Cheshire.

EVENTS



Iran goes nuclear

Self-sustaining nuclear reactions have begun inside Iran's first commercial nuclear power plant. On 10 May, Atomstroyexport, a Russian state-owned firm building the Bushehr nuclear plant (pictured), said it had begun power tests of the 915-megawatt pressurized-water reactor. Construction was begun in 1975 by Siemens but was suspended after the 1979 Islamic Revolution. Russia resumed construction in 1995.

Reactor meltdown

The unit 1 reactor at Japan's Fukushima Daiichi nuclear plant melted down entirely after a massive earthquake and tsunami struck on 11 March, according to analysis from

the plant's owners, the Tokyo Electric Power Company. Data provided by recalibrated equipment inside the reactor indicates that the fuel rods had lost their surrounding coolant four and a half hours after the tsunami arrived. Most of the fuel had probably already fallen to the bottom of the vessel by the time it was flooded with sea water. The full meltdown will complicate future clean-up efforts. See go.nature.com/frm7uk for more.

Shuttle launch

NASA's space shuttle *Endeavour* launched from the Kennedy Space Center in Florida for its final flight on the morning of 16 May. The trip is the penultimate mission of the shuttle fleet. See page 262 for more.

BUSINESS

Chilean dams

Officials in Chile last week approved the construction of five hydropower dams across two major rivers in Patagonia — part of the US\$7-billion HidroAysén project, which aims to generate 2.75 gigawatts of power for Chile. But violent demonstrations followed the approval of the dams, and scientists have criticized the environmental-impact assessment used to justify the scheme. Chilean electricity utilities Colbún and Endesa

COMING UP

22–26 MAY

The American Astronomical Society meets in Boston, Massachusetts.

go.nature.com/q4otmo

22–27 MAY

At its general meeting in Paris, the World Organisation for Animal Health will celebrate the global eradication of rinderpest, a devastating cattle disease.

go.nature.com/z3n2hj

still need permission to build a transmission line that would carry power thousands of kilometres from the remote Aisen region to Santiago. See go.nature.com/gnbvasl for more.

Genomics research

Mount Sinai School of Medicine in New York is setting up a genomics research institute in collaboration with Pacific Biosciences, of Menlo Park, California, which is developing technology for real-time, single-molecule DNA sequencing. The company's chief scientist, the charismatic computational biologist Eric Schadt, will direct the new institute and retain his position with Pacific Biosciences, the Mount Sinai School announced on 16 May.

Hepatitis milestone

On 13 May, the US Food and Drug Administration approved the first drug to directly target the hepatitis C virus. The drug, boceprevir (Victrelis), is made by Merck, based in Whitehouse Station, New Jersey. Another drug, telaprevir (Incivek) — made by Vertex Pharmaceuticals in Cambridge, Massachusetts — is expected to be approved by 23 May.

NATURE.COM

For daily news updates see:

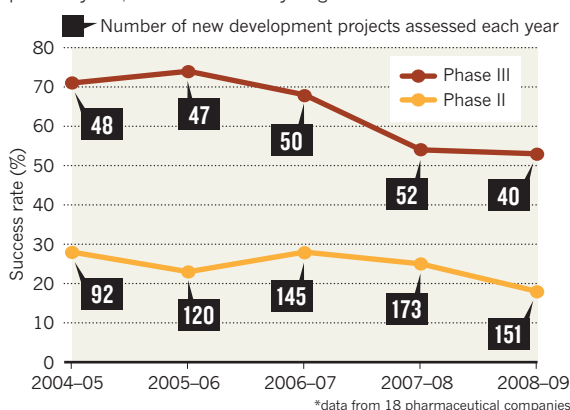
www.nature.com/news

TREND WATCH

The number of drug candidates entered into early clinical trials increased from 2000 to 2009, but that hasn't led to more drugs in the clinic. Success rates for late-stage (phase II and phase III) trials are falling, says the Centre for Medicines Research in London. John Arrowsmith, a scientific director at Thomson Reuters, which owns the centre, says the attrition rate is “unsustainably high” (*Nature Rev. Drug Discovery* **10**, 328; 2011). A clear-out of weak candidates in 2009 may improve future success rates, he adds.

PHARMA'S FALLING SUCCESS RATE

Although more drugs were pushed into clinical trials over the past few years, success rates at key stages declined.



NEWS IN FOCUS

SPACE Final shuttle flights leave NASA facing a void **p.262**

JAPAN Universities take energy savings to an extreme **p.263**

SPAIN A fight to the finish at a top cancer centre **p.264**

CLIMATE California's lonely struggle to cut carbon **p.268**

S. DAS/PANOS



IPCC chairman Rajendra Pachauri faced calls to quit after errors were found in a key report.

CLIMATE CHANGE

Major reform for climate body

Intergovernmental panel aims to become more responsive.

BY QUIRIN SCHIERMEIER

After months of soul-searching, the Intergovernmental Panel on Climate Change (IPCC) has agreed on reforms intended to restore confidence in its integrity and its assessments of climate science.

Created as a United Nations body in 1988 to analyse the latest knowledge about Earth's changing climate, it has worked with thousands of scientists and shared the Nobel Peace Prize in 2007. But its reputation crumbled when its leadership failed to respond effectively to mistakes — including a notorious error about the rate of Himalayan glacier melting — that had slipped into its most recent assessment report (see *Nature* **463**, 276–277; 2010).

That discovery coincided with the furore over leaked e-mails from the University of East

Anglia's Climatic Research Unit in Norwich, UK (see *Nature* **462**, 397; 2009). Some e-mails seemed to show that leading climate scientists, who had contributed key findings to previous IPCC reports, had tried to stifle critics. This put the panel — especially its chairman, Rajendra Pachauri — under intense pressure. The InterAcademy Council, a consortium of national science academies, was commissioned to review the structure and procedures of the IPCC and to suggest improvements to its operations (see *Nature* **467**, 14; 2010).

The council identified the lack of an executive body as a key factor in the IPCC's failure to respond to the crisis. It also urged the panel to improve the transparency of its assessments and to make its

communication and outreach activities more professional. The IPCC adopted several minor changes at a meeting last October (see *Nature* **467**, 891–892; 2010).

More substantial reforms were signed off last week in Abu Dhabi at a meeting of delegates from IPCC member states. An executive committee will be created to oversee the body's daily operations and to act on issues that cannot wait for full plenary meetings. The 13-strong committee will be led by the chairman, and includes the vice-chairs and co-chairs of its working groups and technical support units.

A new conflict-of-interest policy will require all IPCC officials and authors to disclose financial and other interests relevant to their work (Pachauri had been harshly criticized in 2009 for alleged conflicts of interest.) The meeting also adopted a detailed protocol for addressing errors in existing and future IPCC reports, along with guidelines to ensure that descriptions of scientific uncertainties remain consistent across reports. "This is a heartening and encouraging outcome of the review we started one year ago," Pachauri told *Nature*. "It will strengthen the IPCC and help restore public trust in the climate sciences."

The first major test of these changes will be towards the end of this year, with the release of a report assessing whether climate change is increasing the likelihood of extreme weather events. Despite much speculation, there is scant scientific evidence for such a link — particularly between climate warming, storm frequency and economic losses — and the report is expected to spark renewed controversy. "It'll be interesting to see how the IPCC will handle this hot potato where stakes are high but solid peer-reviewed results are few," says Silke Beck, a policy expert at the Helmholtz Centre for Environmental Research in Leipzig, Germany.

The IPCC overhaul is not yet complete. Delegates postponed a decision about the exact terms of office of the group's chairman and head of the secretariat. Critics say that these terms should be strictly limited to the time it takes to produce a single assessment report, about six or seven years. With no clear decision on that issue, Pachauri could theoretically remain in office beyond 2014, when the next full report is due for release.

But the Indian economist says he has not considered staying on that long. "My job is to successfully complete the next assessment," he says. "That's what I'm solely focused on." ■

NATURE.COM
Read more on climate controversy at:
nature.com/climategate



The space shuttle *Endeavour* launched for the last time on 16 May. Only one shuttle flight remains.

SPACE EXPLORATION

Shuttle's end spells change at NASA

As the shuttle flies its penultimate mission, the US space agency seeks to fill a looming gap in crew transport.

BY ALEXANDRA WITZE

Near the Kennedy Space Center in Florida, where the space shuttles thunder into orbit, roadside signs reveal the deep ties that the local community feels to the US space programme. The ties are both spiritual and economic: church signs wish the shuttle *Godspeed* before each launch, and liquor stores tout their selection as 'out of this world'.

That relationship is now heading for an extended and painful hiatus. On 16 May, after several weeks of delays, the space shuttle *Endeavour* embarked on the penultimate shuttle flight, carrying a large cosmic-ray

detector (see *Nature* 473, 13–14; 2011) to the International Space Station (ISS). The final, 135th launch of the 30-year shuttle programme will take place by late summer, when *Atlantis* is set to take flight to ferry another load of astronauts, equipment and supplies to the station.

"Although we're ending the space-shuttle programme, we are not ending the nation's human space-flight programme," says Philip McAlister, acting director of NASA's commercial space-flight development programme. "It's evolving into an exciting new

NATURE.COM
Seven experts
ponder the future of
NASA:
go.nature.com/izs9zr

paradigm." Before that new phase begins, however, NASA will face years with no crewed spaceship of its own, and the Space Coast, the region around the Kennedy centre, will lose some 8,000 space-related jobs.

In the short term, NASA will have to buy seats aboard Russia's Soyuz capsules, now the only way to deliver crew members to the ISS. The next such flight, slated for 7 June, will carry one astronaut each from Russia, Japan and the United States to the station. The United States also has places reserved on the ten further Soyuz flights that are scheduled before the end of 2013.

In the long term, NASA expects to give private companies the responsibility of getting astronauts into low-Earth orbit, an approach championed by US President Barack Obama. On 18 April, the agency announced that it would share US\$269 million among four companies developing commercial space-flight options: SpaceX, of Hawthorne, California; Boeing, of Houston, Texas; Blue Origin of Kent, Washington; and Sierra Nevada of Louisville, Colorado. All say that they will begin flying crewed spacecraft in 2014–15.

SpaceX has already flown its unmanned Dragon capsule into orbit and recovered it successfully, the only private company to manage such a feat. NASA is expected to decide in the coming weeks whether to combine the second and third Dragon flights, scheduled for later this year, to advance directly to a mission that docks with the ISS — an option strongly endorsed by Elon Musk, chief executive of SpaceX.

For NASA, which has always relied on its own or Russian rockets to get astronauts into space, the commercial crew-transportation business represents a fundamental shift to a different — and untested — way of doing things. Many within the agency are upbeat, despite the unknowns.

"The shuttle was one of the ways to go to space," says Chris Hadfield, an astronaut with NASA and the Canadian Space Agency, who will spend six months on the ISS next year, three as commander. "It was not the only way."

Commercial suppliers will also be needed to ferry cargo — including research experiments — to the ISS. The Soyuz flights can carry very little payload, says Tara Ruttley, the station's associate programme scientist. NASA has contracted for 12 cargo flights with SpaceX and 8 with Orbital Sciences of Dulles, Virginia, to begin as early as next year. Russia's unmanned Progress resupply ships can also carry research payloads up to the station.

Freed of the need to develop transport for low-Earth orbit, the thinking goes, NASA can focus on the task handed to it by Congress in last October's NASA Authorization Act. By 2016, the agency is supposed to develop a heavy-lift rocket and a crew vehicle to send astronauts to distant targets such as the Moon or near-Earth asteroids. The rocket will draw

AP PHOTO/J. RAOUX

on work done under the now-defunct Constellation programme, developed during the previous administration as a next-generation replacement for the shuttle.

The new programme reflects the concerns of congressional lawmakers who fear the decline of NASA and of regions dependent on the jobs it has provided, such as the Space Coast (see *Nature* 472, 16; 2011). In an opinion piece published in the *Orlando Sentinel* on 26 April, US Senator Marco Rubio (Republican, Florida),

chastised Obama for not allocating enough money for the heavy-lift vehicle to meet its target launch date in 2016. “The bottom-line impact of the president’s space agenda is a full retreat from America’s long-standing commitment to space exploration,” Rubio wrote. Others echo Rubio’s ire — something that could come back to haunt Obama, given Florida’s probable role as a key battleground state in the 2012 presidential election.

For shuttle workers facing imminent job

loss, commercial flights seem a distant dream. Engineers have been dismantling shuttle-related equipment at Kennedy’s spare launch pad and preparing the already-retired *Discovery* shuttle for shipment to its final home at a Smithsonian Institution museum near Washington DC. For now, the era that is passing commands far more attention than the one that is promised, as a technical community that is used to making space a way of life settles in for a long stint on the ground. ■

ENERGY

Japan rethinks its energy policy

Renewables come to the fore as universities take the lead on electricity conservation.

BY DAVID CYRANOSKI

In the aftermath of the Fukushima disaster, power shortages have forced one of the world’s most energy-efficient countries to make do with even less. That may become the norm after Prime Minister Naoto Kan last week shelved a 2010 goal to build 14 nuclear reactors over the next 20 years.

With Japan’s energy policy in tatters, advocates of renewable energy and efficiency savings are seizing the opportunity to argue their case. But the measures will have to make up a major shortfall: under the previous plan, the country’s nuclear generating capacity was set to double, to meet half of the nation’s electricity needs (see *Nature* 472, 143–144; 2011).

Kan had little choice but to change course after the accident at the Fukushima Daiichi nuclear plant, which was triggered by an enormous earthquake and tsunami on 11 March. Demonstrators have been calling for the closure of some or all of Japan’s 54 nuclear reactors ever since. In addition to ditching plans for new plants, Kan has promised a much greater emphasis on efficiency measures and renewable sources, which Japan has been relatively slow to adopt. “Before 11 March, there was a black cloud over energy policy, formed by industry and the industry ministry. Now there’s a crack in that,” says Tetsunari Iida, executive director of the Institute for Sustainable Energy Policies in Tokyo, which advises the government on renewable energy.

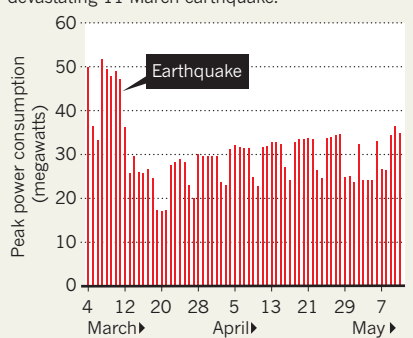
Last month, the institute set out an ambitious vision for the country’s energy mix. Its plan for the Tohoku region calls for energy demand to be reduced such that all the region’s needs can be met by renewable sources by

2020. Because Tohoku saw the greatest damage from the earthquake, and has a lot of potential to exploit wind power, the institute thinks that renewables could be introduced quickly while creating jobs to regenerate the local economy.

Countrywide, the institute says that renewables’ share of the energy mix should rise from about 8% to 30% by 2020, and to 100% by 2050, a strategy that requires demand to be halved. “It

POWER CUT

The University of Tokyo is using at least 30% less electricity than it was just before the devastating 11 March earthquake.



is technically feasible but politically challenging,” says Iida. Iida helped to draft a bill that guarantees utilities a high price for electricity from renewables. The bill was, coincidentally, finalized by the government on 11 March and is expected to become law next month.

Tatsuo Oyama, an engineer at the National Graduate Institute for Policy Studies in Tokyo who models electricity investment scenarios, says that the target might be possible in Tohoku. But he warns that the country should not become dependent on unpredictable energy sources, and that the waning support for nuclear energy could be reversed if Japan tackles the aftermath of the Fukushima disaster effectively.

Tokyo’s universities are already proving that the country still has room for energy savings. The University of Tokyo, for example, has cut peak power usage by 30–40% by turning off lights and air-conditioning, shutting down extra lifts and running energy-intensive experiments at night (see ‘Power cut’).

Researchers at the university say that their low-energy lives are inconvenient, but largely manageable. Restricting the use of some equipment to off-peak hours is “realistic and feasible,” says neurochemist Haruhiko Bito, although he adds that scheduling researchers’ energy use can be time-consuming and depressing. And chemist Eiichi Nakamura says that the loss of instruments and computer systems has slowed research. “The electricity shortage made us realize that we can indeed save energy easily by 10%,” but that 30% cuts will impact productivity in the longer term, he says. Others worry that the strategy will discourage younger scientists by forcing them to work at night.

The challenge will only intensify in the sweltering summer months. Animal facilities and sensitive instruments will take priority for precious cooled air, while professors and students sweat in rooms with minimal air-conditioning. Staggered work schedules and holidays are being considered to mitigate the effects.

With Japan’s energy strategy in flux, these conservation policies will probably be in place for the foreseeable future. “This is tough but, in another sense, this has a positive aspect,” says Toshio Yamagata, an ocean modeller at the University of Tokyo who has had to cope with a 30% cut in supercomputer operating time. “It is a good occasion for us to realize our resources are not infinite.” He adds that a greater appreciation of energy could lead to more elegant experiments. “We will design experiments more carefully and digest the results in a deeper way, rather than just obtaining tonnes of data.” ■



WWW.NATURE.COM/JAPANQUAKE

SOURCE: UNIV. TOKYO



CNIO director Mariano Barbacid (left) and science minister Cristina Garmendia are at loggerheads.

POLICY

‘Soap opera’ sours cancer chief hunt

Spain’s scientists worried by fallout from high-profile spat.

BY ALISON ABBOTT

A bitter row between a prominent Spanish biologist and the country’s science minister has spiralled into a potentially damaging leadership crisis at the country’s top cancer research institute.

The Spanish National Cancer Research Centre (CNIO) in Madrid is trying to find a replacement for molecular oncologist Mariano Barbacid, who became the centre’s founding director in 1998. But an international committee of five high-ranking scientists has now withdrawn from the process after its search for a director became caught up in the quarrel

between Barbacid and the minister for science and innovation, Cristina Garmendia, who has repeatedly clashed with Spanish scientists (see ‘New path for researchers’).

Members of the search committee were dismayed early this month when their shortlist of four candidates was leaked to the press, and then concerned when Garmendia’s ministry scheduled a meeting of the CNIO’s governing board to announce the new director for 16 May, giving them just ten days to make a decision. The top candidate withdrew last week, although this may not be connected with the row. The governing board is now seeking new candidates and deferring further

discussions until 22 June. For researchers at the institute, events are playing out like an unhappy soap opera and could harm their chances of attracting a top player.

Barbacid, a co-discoverer of the first cancer-causing gene, or oncogene, announced in September 2009 that he wished to step down as director to concentrate on his research, once a successor could be found. The trouble began when the ministry could not offer suitable employment conditions for international candidates. A new search started, and there were plans for one of the institute’s vice-directors to become interim director — until the small bombshell of the Experimental Therapeutics Programme dropped in December 2010.

Barbacid had created the programme in 2006, aiming to develop drug candidates based on chemicals that selectively target oncogenes. Using government loans to support the research, he patented several small molecules in October 2010. But at around the same time the ministry declined to extend the loans.

Barbacid says he quickly found two private investors to cover the shortfall, but the ministry rejected this plan in December, saying that its legal counsel had advised that the arrangement was incompatible with the laws governing research foundations. Barbacid was furious. “I checked with good lawyers who made sure the scheme was in compliance with the law,” he says. “Garmendia’s lawyers are just wrong — and her intentions were malicious.” The ministry declined to comment to *Nature* on the row.

On 4 March, the Spanish parliament approved a new law that allows public research foundations to attract private funding for research and innovation. But the ministry did not change its position on the private funding for the oncogene drug programme. When Barbacid issued a press release on 3 May describing a recently published drug target, he took the opportunity to criticize the ministry for preventing him from developing the findings into a potential therapy. The ministry quickly hit back at Barbacid for “raising false hopes in cancer sufferers”, and for seriously breaching professional ethics. It then announced that his successor would be unveiled on 16 May. “It was a visceral response to get rid of me,” says Barbacid. A ministry spokesman denied there was a connection between the two announcements.

Miguel Ángel Piris, a former CNIO vice-president who left in February to take up a position at the University Hospital Marques de Valdecilla in Santander, Spain, says he “hopes that the clash will not harm the prestige of the CNIO or its director, who has managed to create a respected and prestigious institution”. Miguel Beato, director of the Centre for Genomic Regulation in Barcelona, adds that the affair underlines a weakness in Spain’s science system. “Spain doesn’t have research councils or authoritative research agencies who can mediate such rows,” he says. “Politicians have too much direct influence on science here.” ■

SPAIN’S SCIENCE LAW

New path for researchers

After a two-year gestation, a controversial update to Spain’s science law was finally approved by Congress on 12 May. The bill aims to create a research framework akin to those of other European countries, including a structured career path and an independent research-granting agency. The legislation has been a key goal of science minister Cristina Garmendia since she took office in 2008, and researchers have campaigned vigorously to ensure that the final bill protects jobs and funding.

But the final law does not go far enough

for the Dignified Research campaign, supported by 2,500 scientists. They had asked the government to implement five-year, tenure-track contracts with regular evaluations that, if passed, would lead to a guaranteed job. Instead, the bill creates ‘access contracts’ for postdoctoral researchers that do not necessarily lead to a permanent position. The Confederation of Spanish Scientific Societies is also concerned that the law fails to establish clearly the independence of the granting agency. **Michele Catanzaro**



The honeybee is under threat from a formidable array of pathogens, including the Varroa mite seen here.

APIOLOGY

Geneticists bid to build a better bee

Honeybee genome offers clues for fighting diseases.

BY GWYNETH DICKEY ZAKAIB

For Scott Cornman, the honeybee genome is a prized resource, yet he spends much of his time removing it. Cornman, a geneticist for the Bee Research Laboratory of the US Department of Agriculture (USDA) in Beltsville, Maryland, is trying to characterize the various pathogens that plague the honeybee (*Apis mellifera*), arguably the world's most important insect. His strategy is to subtract the honeybee genome from every other stray bit of genetic residue he can find in bee colonies, healthy and diseased. The remaining genetic material gives a complex metagenomic portrait of other organisms that inhabit the bee's world, including viruses, bacteria and fungi — some novel — that, alone or in combination, might push a bee colony into precipitous decline.

"Right now we're in the discovery phase, where we're trying to identify what's present,"

says Cornman. "Then we can start looking at the interactions of pathogens and see if they're more virulent than any by themselves."

Cornman was among 100 or so researchers in attendance last week at the Honey Bee Genomics & Biology meeting, held at Cold Spring Harbor Laboratory in New York. It was the first dedicated conference on the topic since researchers met four years ago, soon after the honeybee genome was sequenced (Honeybee Genome Sequencing Consortium *Nature* **443**, 931–949; 2006), and for many it was a chance to marvel at a field transformed.

"There has been a lot of progress made on how disease affects honeybees at the molecular level," says Christina Grozinger, director of Pennsylvania State University's Center for Pollinator Research in University Park, one of the conference organizers. Around the same time that the genome was first published, honeybee colonies across much of the Northern

Hemisphere began to show alarming declines. A syndrome dubbed colony collapse disorder (CCD) has been causing the insects to die off in large numbers, leaving well-provisioned hives suddenly empty. Meanwhile, other parasites, such as the Varroa mite (*Varroa destructor*), which spreads harmful viruses, continue to take their toll. Annual surveys in the United States show that almost 35% of all colonies die during a typical winter. Genomics is yielding new clues to the still-mysterious phenomenon, as well as potential strategies for protecting the insects from a multitude of threats.

At the meeting, Cornman presented data showing that hives affected by CCD have higher levels of microscopic gut fungi called *Nosema*, and a greater prevalence of several viruses, two of which had not been detected in bees before.

Yet despite having a multitude of enemies, many bees are holding their own, says research entomologist Jay Evans of the USDA's bee laboratory. "The question is not why are bees getting sick, but how are they surviving against this onslaught of parasites," he says.

The genome offers a window into the bees' immune pathways, Evans adds. The goal is to identify the genes that are crucial in helping bees thwart attack, and, ultimately, to strengthen these defences. "You can breed for these traits, but with genetic markers you could do it faster," he says.

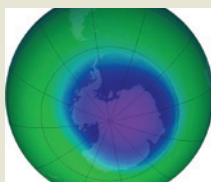
In cases in which nature cannot do the job, some researchers are now exploring more direct ways of boosting bees' resilience. In some insects, double-stranded RNA, a hallmark of viral infection, can provoke a specific antiviral immune response. At the meeting, Michelle Flenniken, a virologist at the University of California, San Francisco, presented evidence that, in honeybees, it can also trigger a general immune response that might ward off a variety of threats. "This may be a new viral response that hasn't been well-characterized in honeybees," says Flenniken, who is exploring the genes involved in the process. "What we think we've found is a window into this new immune-response pathway."

Flenniken adds that knowing more about the bee's immune responses might help researchers to find ways of "priming the system" and help bees to cope with their foes at the genomic level. Such a prospect may be a long way off, but it's certain to keep researchers abuzz until their next gathering. ■



**MORE
ONLINE**

TOP NEWS



The Antarctic ozone hole is beginning to recover
go.nature.com/cwqbed

OTHER NEWS

- An economic case for marine protected areas go.nature.com/jzpktu
- Bacteria helped early animals to breathe go.nature.com/zlpnzd
- Missing University of Texas physicist may be in a jail in Iran go.nature.com/rat6v9

EXPLAINER



Why the Mississippi floods should have been expected
go.nature.com/sr1zjc

NUCLEAR ENERGY

Battle of Yucca Mountain rages on

Proposed interim storage unlikely to settle US debate.

BY JEFF TOLLEFSON

Staff have been cut, contractors laid off, offices closed and even furniture disposed of. But despite all its efforts to back away from plans to store spent nuclear fuel deep under Yucca Mountain, Nevada, the administration of US President Barack Obama just can't seem to bury the idea.

An expert commission appointed by the administration is looking for an alternative solution. On 13 May, at a public meeting in Washington DC, commissioners discussed some preliminary recommendations: create one or more centralized facilities at which waste would be temporarily stored in dry casks, while engaging with the public in a new process to identify a permanent repository for the piles of spent nuclear fuel accumulating at US reactors. But given the history of doubts about the site's geology and the state-wide opposition that has plagued Yucca Mountain since it was singled out by the US Congress nearly a quarter of a century ago, many are sceptical that a more palatable answer will emerge.

"It is important that there will be a consensus recommendation, but it is our view that most of the issues associated with used nuclear fuel have been considered for a long time," says Alex Flint, senior vice-president for governmental affairs at the Nuclear Energy Institute in Washington DC. The ongoing nuclear disaster in Japan is adding some urgency to the question, Flint says, but "it hasn't made reaching agreement any easier".

Meanwhile, an 8 April report from the Government Accountability Office (GAO), an independent arm of Congress, says the Department of Energy should "develop a preliminary plan to restart the project" at Yucca Mountain, anticipating that future policy shifts — and a pair of legal challenges from states that want to get rid of the waste piling up within

their borders — may force it to do just that.

Even after decades' worth of research on the site, costing more than US\$15 billion, doubts remain over the technical suitability of Yucca Mountain, given factors such as seismic activity and water infiltration. But the politics were crystal clear when Obama promised to shut it down during his 2008 election campaign. In keeping with that promise, last year the energy department filed a motion to withdraw its application to store nuclear waste at Yucca Mountain — offering no technical or scientific reasons for the reversal, except to say that the project was "not a workable option".

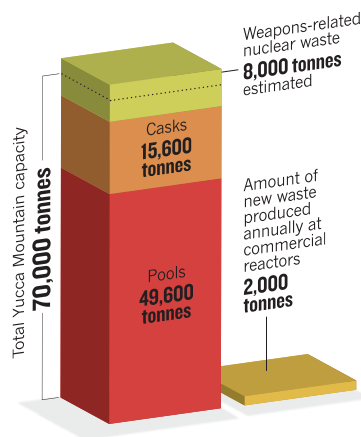
By then, however, the Nuclear Regulatory Commission (NRC) was in the midst of a regulatory assessment that — barring the inevitable lawsuits — could have cleared the way for waste shipments to Yucca Mountain, as directed under a federal law signed by President George W. Bush in 2002. Now the department's decision to withdraw is being challenged within the NRC, in federal courts and on Capitol Hill.

"Now that this administration has decided to ignore the law, our nation has no long-term storage plans for radioactive wastes," lamented Republican (Georgia) representative Paul Broun, chairman of the House Science Subcommittee on Investigations and Oversight, during a congressional hearing last week. He and other Republicans on the science panel, together with two other House committees, are challenging the administration's decision and demanding documentation.

Their complaints have new resonance in the wake of Japan's Fukushima Daiichi power plant disaster, in which radioactivity from nuclear waste stored at the plant apparently escaped into the environment. In the United States, more than 65,000 tonnes of spent nuclear fuel from commercial reactors currently sit in temporary storage, with around 2,000 more

A GROWING DILEMMA

Nuclear waste in temporary storage in the United States already exceeds the limit set for Yucca Mountain.



tonnes accumulating every year (see 'A growing dilemma'). Combined with waste from weapons programmes, the amount surpasses what has been set as Yucca Mountain's statutory limit, although there is room to expand should the site find itself back in business. Much of the waste resides in storage pools at reactor sites, like those at Fukushima.

Additional pressure is coming from states that want their spent nuclear fuel moved out. Washington and South Carolina are leading challenges to the Department of Energy's decision to withdraw from Yucca Mountain, both at the NRC and in the District of Columbia Federal Appeals Court. "Our reading of the law is that the issue needs to be concluded on the basis of its technical merits," says Mary Sue Wilson, a lawyer working on the case for the Washington State Attorney General.

Within the NRC itself, the Atomic Safety and Licensing Board ruled last year that the energy department does not have the legal authority to withdraw its application. A final decision on that case is now pending before the full commission, which is chaired by Gregory Jaczko, a political appointee. He is the former

chief of staff to Senate majority leader Harry Reid of Nevada, who has spearheaded opposition to Yucca Mountain, and

many believe that Jaczko is stalling to prevent a ruling against the administration.

"The chairman controls when the NRC votes, and the chairman doesn't like the current vote," says Lake Barrett, a consultant and former deputy director at the energy department. NRC officials say the commissioners are still deliberating on the issue.

Speaking at the Annual Nuclear Industry Conference and Nuclear Supplier Expo in Washington DC on 11 May, deputy energy secretary Daniel Poneman told industry officials that the administration is hoping the presidential commission will find a way to reshape the discussion and build the kind of consensus that will at last allow the country to move forward.

"Clearly, the mistake we made in 1987 was jamming it down the throat of the Nevadans," says Phil Sharp, a commission member and president of Resources for the Future, a think tank based in Washington DC. Sharp says the government must work with the public and communities, presenting nuclear waste disposal as a national priority in a way that appeals to people's patriotism.

"Yucca Mountain has nine lives, and nobody knows how many lives have been used up."

The commission intends to issue a draft report in July and a final one next January. With its recommendations in hand, the administration is expected to propose legislation that would establish a new process for identifying nuclear waste storage sites.

Yet such a process could well take decades, the GAO report concludes, and the government's reversal at Yucca Mountain could serve to galvanize public opposition at other candidate sites. Since the debate began, "no states have expressed an interest in hosting a permanent repository for this spent nuclear fuel ... including the states with sites currently storing the waste", the report adds. The commission's scheme for an interim storage facility may prove no more appealing, given fears that 'interim' means permanent as long as the present impasse continues. Such fears have in the past halted interim storage proposals in states such as Wyoming. And even if one community decides that it is willing to play host to the waste, that doesn't mean others won't challenge nuclear-waste transportation routes.

Nevertheless, the nation will need to find a permanent repository at some point, and Yucca Mountain, it seems, is down but not out. "Yucca Mountain has nine lives," says Ed Davis, a nuclear consultant who heads the Pegasus Group in Washington DC. "And nobody knows how many lives have been used up." ■

➔ NATURE.COM
Does US nuclear
power have a future?
go.nature.com/iqisuj



America's top CLIMATE COP

The United States has abandoned comprehensive greenhouse-gas curbs, but California is pressing ahead. Mary Nichols is leading the fight against emissions.

BY JEFF TOLLEFSON

Mary Nichols can take some pride in the view as she travels out of Los Angeles. The San Gabriel Mountains rise up to the north, framed by blue sky with just a touch of midday haze. The clear vista comes in large part because of the California Air Resources Board (CARB), the agency that Nichols leads, which has spent decades cleaning up the city's air. Now she and her team are setting their sights even higher — with an ambitious plan to cut California's greenhouse-gas emissions.

With an economy that outranks all but eight countries, California is a political and economic heavyweight that has never been afraid to flex its muscles. It is big enough to make an impact, and now that politicians in Washington DC have abandoned attempts to enact a national climate law, California is forging ahead on its own. Nichols feels the burden of that strategy acutely, and she is well aware of the challenges ahead.

In the run-up to the state elections last November, many feared that Californian voters would follow Washington DC's lead and cast aside the state's landmark climate legislation,

AB 32. The 2006 law requires a 10% reduction in greenhouse-gas emissions by 2020, and critics — fuelled in part by donations from the fossil-fuel industry — argued that the state's economy was too fragile to withstand aggressive new regulations. But voters turned out en masse to preserve the initiative, which is the first comprehensive climate programme in the United States. California has committed to reducing emissions by the same percentage as the European Union, and the state's unique plan could chart new ground internationally.

Since the 1970s, California has pushed the boundaries of environmental regulation, acting out of both pride and self-preservation. The state has pioneered environmental laws targeting air pollution, water contamination and toxic chemicals. It has advanced the sciences of atmospheric physics and chemistry, developed pollution-control technologies and bullied powerful industries into submission in an epic battle against choking smog in the Los Angeles basin. Other states, and eventually the nation, have followed California's path in

L. ATKINS/SAN FRANCISCO CHRONICLE/CORBIS

developing regulations to control pollution.

But Nichols and her staff at CARB need to go even further to rein in greenhouse-gas emissions. The agency plans to clean up vehicle fuels, promote renewable energy and squeeze more reductions by improving energy efficiency. It is also designing the world's most comprehensive carbon market, set to launch at the start of 2012. Nichols believes that California will one day be able to demonstrate to the rest of the country how environmental protection and economic growth can coexist.

"People in this state are bullish on the ability of California to survive and change, and they fundamentally care about air pollution and environmental issues," says Nichols. "What we do here matters."

From Washington DC to Brussels and Beijing, government leaders will monitor the state's progress closely. Henry Derwent, president of the International Emissions Trading Association based in Geneva, Switzerland, says that California's plans are reassuring governments around the world that all is not lost in the United States. "The overriding feeling in Europe at the government level is relief," says Derwent. "Even though it's not the entire United States, it's a pretty big consolation prize."

CHARM OFFENSIVE

On this day in February, Nichols is travelling from her office in Los Angeles to a conference on sustainable growth at the California State Polytechnic University in Pomona. But the route along Interstate 10 illustrates the scale of the problem. The greater Los Angeles urban area sprawls outwards through towns and cities, filled with millions of people who love their vehicles.

Despite that, Los Angeles has managed to clean its air through a productive interplay between technology and environmental policy. Nichols says that modern vehicles produce 1% of the toxic pollutants emitted by their forerunners in 1975. The city's population has doubled since then and the use of vehicles has grown at an even faster rate, yet the air just keeps getting cleaner.

But CARB now faces a bigger and broader challenge. If no action is taken, California's emissions are projected to climb from 474 million metric tonnes of carbon dioxide equivalent in 2008 to 596 million metric tonnes in 2020. To reach the target set in AB 32, Nichols and CARB must get the total down to 427 million metric tonnes, the amount that the state was emitting in 1990 (see 'Cleaning up California'). To do that, they need to make emissions reductions everywhere they can, and that is what brings Nichols to Pomona.

She is addressing the small conference regarding one of the latest tools in CARB's belt: SB 375, a 2008 law requiring the agency to set targets for greenhouse-gas emissions from vehicles in all metropolitan areas. Her team set those targets last September, and the local and regional planning organizations must now develop strategies to meet them by, for example, promoting public transport, bike lanes and mixed-use zoning that brings amenities to people instead of forcing them to drive.

CARB set a 13%-reduction target for the area that includes Los Angeles, but many local officials complained that the state was imposing costly rules without providing any money to help them comply. Nichols knows that some of those officials are in the audience, and she has come in peace. As she steps up to the microphone, she gives a confident smile and disarms the sceptical leaders by acknowledging that the law's future is in their hands. "You could probably ignore it," she says, scanning the quiet audience for a reaction. "Nothing will happen, as far as I can tell."

Nichols then launches into a pep talk. SB 375 is not a top-down state solution, she says, but a bottom-up tool to help local and regional governments make their communities into more livable places, where people walk and exercise and spend more time with their families and less time alone in cars. This kind of master planning, she says, could set the stage for more organized — and less contentious — development because all parties would have agreed on the basic framework for growth. Nichols then ties up her talk by offering a small cash sweetener, in the form of grants to help local governments get the process started. There

isn't so much as a peep of protest, and by the time lunch rolls around conversations are focusing on how to implement the law.

"If we see there is rising opposition, then we need to act and explain or make adjustments," says Nichols on her way back to the office. That kind of flexibility makes it easier for states than the federal government to negotiate difficult new regulations, she adds. "We are closer to the people that we regulate."

If Nichols makes it look easy, she has had a lot of practice. An environmental lawyer by training, Nichols is a diehard Democrat who has burnished her credentials working for environmental groups. She has also honed her diplomatic skills in various government posts, including a previous stint as head of CARB, from 1979 to 1983. She eventually rose to assistant administrator of air and radiation at the US Environmental Protection Agency (EPA) in 1993, under President Bill Clinton. For Nichols, these political appointments always represented an opportunity to put ideas into practice and put her stamp on the world.

"PART OF BEING A SUCCESSFUL LEADER IS HAVING FOLLOWERS."

By the time the California legislature enacted AB 32 in 2006, Nichols was ensconced in academia as director of the Institute of the Environment at the University of California, Los Angeles. She wasn't looking for a job when Republican governor Arnold Schwarzenegger asked her in 2007 to take over CARB and find a way to meet the target, nor was she particularly thrilled about going to work for a Republican film star. She jokes that when she met Schwarzenegger, she interviewed him for the job, and he passed the test. Convinced that he was genuinely interested in making the programme work, Nichols jumped back into government.

IN THE DRIVER'S SEAT

CARB's plan bets heavily on innovation, some of which the agency is developing and testing at its own facilities. Nichols spends much of her time working from CARB's main science laboratory in El Monte, east of central Los Angeles. This is where agency engineers invented the check-engine light in the 1980s to alert drivers to problems with their vehicle's pollution-control systems. CARB is now developing automated sensors that will allow technicians to more accurately track emissions data in cars using a secure onboard computer. Engineers are busy analysing emissions from advanced vehicles, testing the performance of hybrid electric cars and studying how various technologies could help the state to meet its 2020 goal and a further, non-binding commitment to reduce greenhouse-gas emissions by some 80% by mid-century.

The most ambitious element of CARB's plan is an overarching cap-and-trade programme that will cover roughly 85% of the state's emissions by 2015. Under that system, the state will issue a set number of allowances — initially for free but later through an auction — that companies will need to cover their greenhouse-gas emissions. The total number of permits will decrease each year, and companies will need to either reduce their emissions or buy spare allowances from other companies that have made reductions more cheaply.

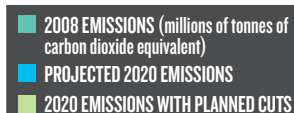
The cap-and-trade programme is an insurance policy. On their own, individual regulations for vehicle efficiency, renewable energy and other items will lower emissions, but they do not guarantee that the state will meet its targeted reductions. The cap-and-trade programme should — if it ever gets off the ground. In March, a California judge determined in a preliminary finding that CARB had failed to do a proper environmental analysis of the programme. The agency is now awaiting a final ruling on how to proceed, but CARB officials hope that the programme will move forward on schedule to begin next year.

Meanwhile, the agency is pressing ahead with other bold plans. CARB is working with partners in Brazil and Mexico to design what would be the world's first market-based programme to allow businesses to offset their emissions by protecting tropical forests. The agency is also establishing another type of offset, involving ozone-depleting compounds

➔ NATURE.COM

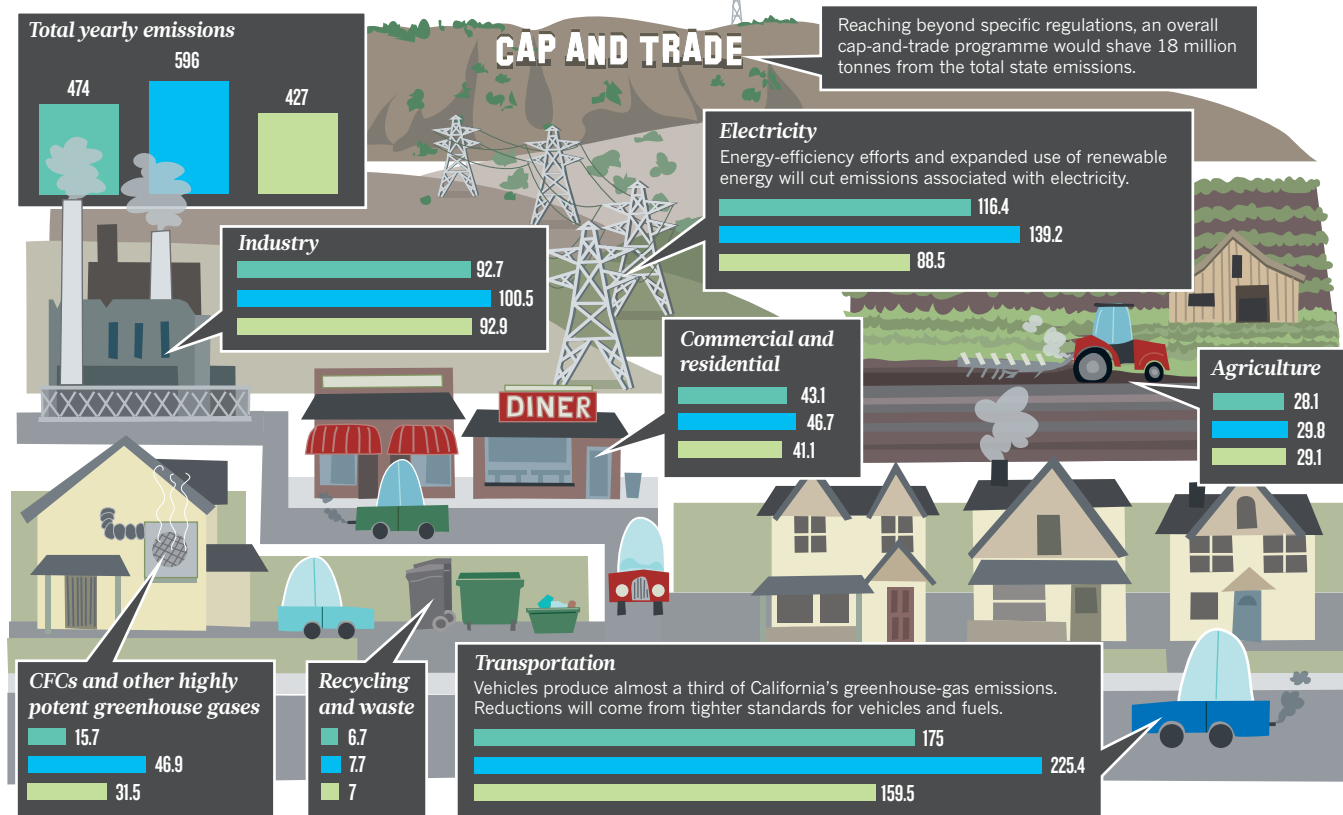
For more on the future of energy, visit:

go.nature.com/jyfh2i



CLEANING UP CALIFORNIA

A state law mandates that California must shrink its greenhouse-gas emissions by 28% from the levels currently projected for 2020. It plans to do this through regulations targeting individual sectors, in combination with an overarching cap-and-trade programme, which imposes a cost on emissions.



such as chlorofluorocarbons, which are powerful greenhouse gases not included in the United Nations' 1997 Kyoto Protocol for reducing greenhouse-gas concentrations. Companies in California could avoid reducing their emissions of carbon dioxide or other Kyoto gases by curbing — or paying someone else to curb — their emissions of the non-Kyoto greenhouse gases, which are not targeted by a more limited cap-and-trade scheme launched by the European Union in 2005.

CARB is trying to avoid pitfalls revealed by the European programme. That scheme, for example, initially issued too many allocations, which led to a collapse in the price of carbon. CARB is taking care to keep an inventory of emissions, so that it can issue an accurate number of initial allowances. But the inventory is calculated in part from figures provided by polluters, so CARB is also carrying out an independent check, funding scientists to measure concentrations of greenhouse gases and other pollutants in the field and then calculate emissions from that data. Already, CARB knows that methane emissions around Los Angeles are higher than the inventory suggests.

Nicholas Bianco, a senior associate at the World Resources Institute in Washington DC who advises agencies on emissions reduction, says that the California cap-and-trade scheme represents a major step forward. "It will be the first of its kind in the world."

James Sweeney, director of the Precourt Energy Efficiency Center at Stanford University in California, says that what is happening in the state is exciting, but he has two fears. The first is that funding for energy and climate research will dry up in the current budget crisis, making the challenge of meeting long-term greenhouse-gas reduction targets in California and elsewhere even more difficult. The second relates to scale. California is important, but it represents just 7% of US emissions.

"The bottom line," says Sweeney, "is that if California is going to have a real impact it will be as the laboratory for the nation."

The chances of that happening are unclear. Northeastern states have a limited cap-and-trade programme for power plants, but western states have backed away from joining California's scheme — although at least

three Canadian provinces are expressing interest. California isn't big enough to run its own system forever, says Nichols, but the state will stay the course for now. She points out that the agency has history on its side.

When CARB published its first greenhouse-gas regulations for cars in 2004, it quickly ran into legal battles with the automobile industry and the administration of President George W. Bush. But last year, the Obama administration brought the various players together in a deal that essentially established CARB's vehicle regulations as national ones.

"When we started the first round of greenhouse-gas standards, the automobile companies wouldn't even talk to us," says Paul Hughes, who headed the effort as manager of the Low Emission Vehicle programme. Today, Hughes says, car makers are engaged at every step in the process as CARB and the EPA prepare to release identical new standards for California and the nation for model years 2017–25. Due late this year, those regulations are expected to translate into an average fuel-efficiency rating of 20–26 kilometres per litre for cars and trucks — a big jump from the current standard of less than 12 kilometres per litre. For Hughes, it is just a matter of time before other CARB policies diffuse outward and upward into the national scene.

On the drive back from Pomona, Nichols ponders the roller-coaster progress of the past few years. With Obama in the White House, it looked as if the United States was finally gearing up for a serious push on global warming. Then lawmakers rejected the idea, leaving California on its own.

The optimist in Nichols thinks that the United States will eventually find its way on climate. But she is also a realist and has a simple message for the rest of the country. "California set itself up to be at the head of what we thought was going to be a parade, but part of being a successful leader is having followers," she says. "At the end of the day, Californians are not going to accept a lonely role as the sole state in the union that is doing anything in terms of carbon." ■

Jeff Tollefson covers energy and environment for Nature in Washington DC.

The growing pains of pluripotency

The field of induced pluripotent stem cells has grown up fast. Now it is entering the difficult stage.

BY ERIKA CHECK HAYDEN

Five years is an eye-blink in science, but that's all the time it has taken for the concept of adult-cell reprogramming to revolutionize the field of regenerative medicine. In August 2006, Shinya Yamanaka of Kyoto University in Japan told the world that he had turned mouse skin cells into induced pluripotent stem (iPS) cells, capable of becoming many types of cell¹. The following year, he repeated the feat for human cells².

Like human embryonic stem cells, iPS cells could potentially be used as therapies, disease models or in drug screening. And iPS cells have clear advantages: they can be made from adult cells, avoiding the contentious need for a human embryo, and they can be derived from people with diseases to create models or even therapies based on a person's genetic make-up. Scientists predicted that iPS cells would change the face of biology and medicine — and some would say they already have. In the past year or so, researchers have published cellular models derived from iPS cells for a staggering array of conditions, from heart defects³ to schizophrenia⁴. And treatments based on iPS cells are moving toward the clinic: in California, for example, a team hopes to gain approval within the next three years to start treating people with the devastating skin disease epidermolysis bullosa using skin tissue grown from their iPS cells.

Yet work in the past few months has highlighted several potential roadblocks. Reprogramming can be inefficient and induce mutations; the reprogrammed cells cannot develop into some cell types; and those they can generate are not always a good model for disease. New issues are emerging apace: work published last week⁵ shows that, in a particular strain of mice, iPS cells cause immune reactions when they are transplanted into other mice with the same genetic make-up — raising questions about whether it will be possible to transplant iPS-cell-derived tissue back into the person from which it is made. No one doubts that iPS cells still have enormous potential, but the field's initial optimism has cooled. "Right now, we are a long way from being sophisticated enough to take advantage of these cells' potential," says neuroscientist Arnold Kriegstein of the University of California, San Francisco (UCSF). "Things are still at a very early stage." Here, *Nature* looks at some of the field's biggest challenges, and how they are being tackled.



1 FINDING A RECIPE

From the start, biologists have tried to devise safer and more efficient recipes for making iPS cells than Yamanaka's method, which used a retrovirus to deliver a powerful shot of four genetic reprogramming factors into cells. Retroviruses integrate into a host cell's DNA and can therefore potentially disrupt gene expression and lead to cancer; and one of the reprogramming factors, *Myc*, is itself an oncogene that could cause cancer.

To outsiders at least, a new, 'improved' reprogramming method seems to be published every month. But Yamanaka's retroviral method is still the most efficient, and the one used most widely. The retroviral technique can transform about 0.01% of human skin stem cells into pluripotent cell lines; by comparison, adenoviruses, which do not integrate into the genome, transform just 0.0001–0.0018% of cells⁶, and delivering the reprogramming factors directly into a human cell transforms 0.001% (ref. 7). Inefficiency increases the cost and difficulty of deriving iPS cells for cell banks, and poses a particular problem when working with rare cell sources. Researchers have also tried omitting *Myc*, as well as silencing it or stripping it from the cell once reprogramming is complete. But these workarounds also lower the efficiency of reprogramming, and a silenced *Myc* might be reactivated.

Addressing these concerns is already a top priority for the field. Researchers continue to tinker with their reprogramming recipes, trying to find the factors, and means of delivery, that are the most efficient and don't increase the risk of cancer. In April, a group led by Edward Morrissey at the University of Pennsylvania in Philadelphia reported that it could boost the efficiency of reprogramming by two orders of magnitude over standard techniques by using a retrovirus to shuttle in a cluster of microRNAs⁸. "It is very important for us to get these reprogramming methods to work well enough so that we can compare them and see whether they make any difference in the stability of the cells and in tumorigenicity," says developmental neurobiologist Jeanne Loring of The Scripps Research Institute in La Jolla, California. "No one has done that yet, and it is going to be a long haul before we figure this out."

► NATURE.COM
Read more about
iPS cells at
go.nature.com/otnzrl

ILLUSTRATIONS BY THOMAS POROSTOCKY



2 PATCHING THE SCARS

A whole new set of questions has arisen over the past year, concerning the genetic impact of the reprogramming process. In July 2010, groups led by George Daley⁹ at the Children's Hospital Boston and by Konrad Hochedlinger¹⁰ at the Massachusetts General Hospital in Cambridge published studies showing that iPS cells carried an 'epigenetic memory' — chemical modifications in their DNA that had come from the original adult cells and had not been erased by the reprogramming. This, they say, explains why iPS cells cannot generate as many adult cell types as embryonic stem cells can.

Researchers were soon reporting that iPS cells were more likely to contain mutations than cultured human embryonic stem cells. Four groups scoured the genomes of iPS cells for changes in single DNA bases¹¹, DNA rearrangements called copy-number variations^{12,13} and differences in chromosome number¹⁴. The studies found higher levels of all three. Worse, the mutations in iPS cells were not just inherited from the parent cells — some seemed to result from the reprogramming and culture process. Loring's group reported¹², for instance, that a protocol for differentiating iPS cells into cardiac cells selected for cells with genetic rearrangements.

This picture is still coming together. One of the studies¹³ on copy-number variations found that many of the rearrangements disappeared after the iPS cells were cultured over long periods of time, probably because the most severely mutated cells were outcompeted by the genetically healthier ones. But this February, a team led by Joseph Ecker at the Salk Institute for Biological Studies in La Jolla, California, reported that it had detected epigenetic signatures of the parent cells in human iPS cell lines even after they had been cultured many times and differentiated into specific cell types¹⁵. A third study¹⁶ suggested that iPS cells are no worse than embryonic stem cells in this regard. Developmental biologist Alexander Meissner of Harvard University in Cambridge and his team reported that epigenetic and genetic variation was similar across 20 human embryonic stem cell lines and 12 iPS cell lines. "What we see is not so much a lot of variation across iPS cells, but a lot of variation across pluripotent cells," Meissner says.

Researchers expressed their concerns about such effects at a meeting convened by the US National Institutes of Health and the Food and Drug Administration (FDA) in Bethesda, Maryland, on 21–22 March that focused on hurdles to translating research on pluripotent cells into the clinic. The concern is that the mutations could have unpredictable and undesirable effects on the cells, and on the patients they end up in. "The genomic changes are going to be a big deal to the FDA," Loring says.

Meissner's group has devised a 'scorecard' of gene expression and methylation — a type of epigenetic mark — that correlates with an iPS cell line's level of pluripotency. It should help researchers to identify and avoid the practices that generate the worst genetic aberrations, and to screen for the lines that are least affected. And researchers are beginning to examine whether and how these genetic and epigenetic effects affect the capabilities and characteristics of iPS cells.

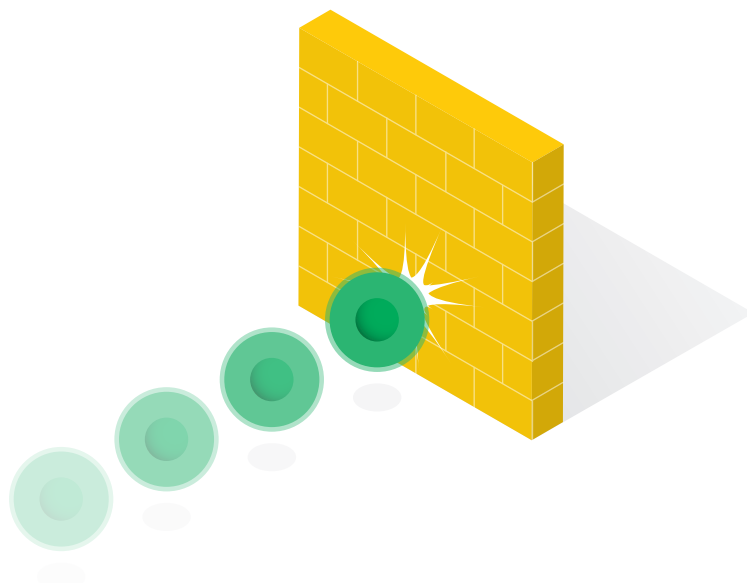
"Right now there are two schools of thought on this," Loring says. "One is that the sky is falling, and the other is that it's a good thing that we're finding out about this now, so that we can discover whether these are biologically relevant changes."

3 HITTING THE LIMITS

iPS cells are immensely flexible, but they can't do everything. Liver cells derived from iPS cells could in theory replace animals in drug toxicology screening, for example. But researchers have struggled to get any human stem cells to differentiate into tissues, such as liver, that are normally derived from the endoderm — the most interior of the three germ layers that make up the embryo. This might be because the series of signals necessary for these cells' development and complex function are difficult to recapitulate. Liver specialist Holger Willenbring at the UCSF points out that hepatocytes have many roles, from detoxifying the blood to making circulating proteins. "There are a lot of different jobs that the cell has to accomplish, and it is hard to get that right in cell culture," he says.

One of the hottest areas of research on human stem cells is in cell-replacement therapy for type-1 diabetes, a disease that develops when insulin-producing cells in the pancreas are destroyed. But no one has been able to make a fully functional and mature insulin-producing pancreatic beta cell — also derived from the endoderm — because researchers don't know the exact series of growth signals necessary and, perhaps, because beta cells usually develop in a three-dimensional environment that is difficult to replicate in a culture dish. Maybe this won't matter, says developmental biologist Matthias Hebrok at the UCSF. Beta-cell 'progenitors' have been made, from embryonic stem cells and from iPS cells, that can secrete insulin. They are less efficient than a normal beta cell, but perhaps efficient enough to help a person with diabetes.

Researchers are making a concerted effort to devise the correct recipe for making mature beta cells and hepatocytes, and Willenbring and others say that those in the field are working as a team on this for the first time. A paper published last week¹⁷ circumvented iPS cells altogether, describing the generation of hepatocyte-like cells directly from mouse skin cells, using a cocktail of regulatory proteins important for liver development. But Willenbring says he still has questions about whether the cells were able to perform all the functions of hepatocytes.



4 MAINTAINING STANDARDS

The relative ease of reprogramming has thrown the iPS cell field open to almost anyone. But from the start, researchers have worried that the low barrier to entry and the extremely competitive pace have meant that standards are not as rigorous as they should be. “There’s a tremendous amount of pressure to get these papers out there, and in the rush investigators are not characterizing their cells very well,” Kriegstein says.

Kriegstein points to a paper published last November¹⁸ by a team led by Alysson Muotri of the University of California, San Diego. The team studied people with a mutated gene that causes the neurological condition Rett syndrome as a model for an autism-spectrum disorder, which causes behavioural difficulties. The researchers derived iPS cells from these patients, and showed that as the iPS cells differentiated into neurons, they initially expressed genes typically found in neural ‘precursor’ cells, then later expressed genes involved in neuronal signalling. The neurons from the patients with Rett syndrome were smaller than those derived from people without the disease, and they also had signalling defects and other differences. But Kriegstein says that this molecular characterization is not enough to show exactly what type of neurons had formed, what part of the central nervous system they represented and how they therefore relate to processes that go awry in the brain. “Possibly the reported abnormalities are relevant to the intended diseases,” Kriegstein says, referring to this and other iPS papers, but this would be very surprising given that most neurodevelopmental and neurodegenerative diseases affect specific populations of neurons at specific times during development, he says.

Muotri says that his team’s analysis of the neurons was thorough, and included electrophysiological tests showing that the cells could fire action potentials and were therefore functional. He also points out that almost all neurons in the brain express the gene, so he did not want to limit the studies just to regions that had been associated with Rett syndrome, as insights from the cultured neurons could illuminate the molecular mechanisms that disrupt the circuits controlling behaviour. “When we find such a cellular phenotype in culture, we know we can now start from there to understand other layers of complexity,” Muotri says.

5 MODELLING THE MIND IS HARD

Investigators are creating patient-specific iPS cells to model almost every disease known — but in some cases, researchers question how much can be learned from the models. The biggest debate is over models of complex neuropsychiatric and behavioural disorders: can reprogrammed cells really mimic conditions such as schizophrenia or autism, which affect the brain and behaviour in complicated ways? “I’ve had clinicians ask if we can make iPS cells from a patient who was mentally retarded,” says developmental biologist Christine Mummery of the Leiden University Medical Center in the Netherlands. But she questions how useful that would be. “I said, ‘I don’t know, how you would measure the IQ of a neuron in a dish?’”

Researchers working with these models argue that they are still valuable. In April, Fred Gage at the Salk Institute and his team reported⁴ that they had derived neurons from the skin cells of a person with schizophrenia, and that some differences between those neurons and normal neurons could be corrected by administration of the antipsychotic drug loxapine. Like Muotri, Gage says that the model is designed to uncover how the genetic factors underlying schizophrenia affect the function of neurons. “Although we cannot measure the behaviour of the patients, we propose that we can measure the activity of the neurons, and the goal is to search for cellular and molecular processes that underlie

the behavioural phenotypes,” he says.

A related problem arises with diseases of ageing, a hot field in iPS-cell research: many of the conditions strike mature cells, so iPS cells — which are essentially starting their developmental lives afresh — might not be relevant. With a disease such as Parkinson’s, says Mummery, “this is a real issue — will you be able to get neurons mature enough to see anything? People are working very hard not only to make their cell type of interest but also to make them mature, so that’s still a major technical obstacle.”

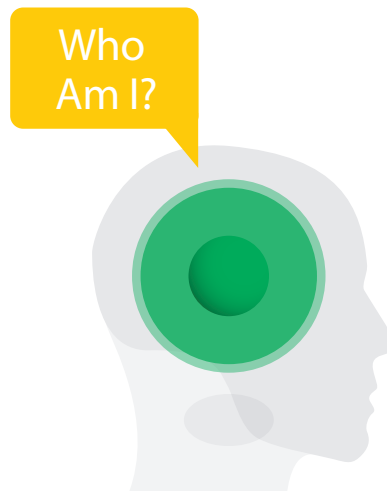
Some researchers counter that even ‘young’ cells show traits related to diseases of ageing. Renee Reijo Pera at Stanford University in California made iPS cells from Genia Brin — the mother of Google co-founder Sergey Brin. She has Parkinson’s, a condition that is marked by the destruction of dopamine-producing neurons. Once differentiated into neurons, the cells secreted dopamine and were more sensitive to chemicals that can induce cell death than were dopamine-secreting neurons derived from healthy people¹⁹. “This seems to be to be the best model of Parkinson’s disease,” Reijo Pera says.

Many iPS-cell researchers see the field’s growing pains as signs that it is reaching a state of maturity; they say that the problems are no different from those that many biomedical research fields face as they inch towards clinical application. “There was this huge euphoria in the beginning, with everyone thinking iPS will do everything, cure all diseases, and be super-easy,” Meissner says.

“But not everyone can become a stem-cell biologist overnight,” he says. “It’s a bit of a reality check that things are not as simple as we thought.” ■

Erika Check Hayden is a senior reporter for Nature based in San Francisco.

1. Takahashi, K. & Yamanaka, S. *Cell* **126**, 663–676 (2006).
2. Takahashi, K. *et al.* *Cell* **131**, 861–872 (2007).
3. Itzhaki, I. *et al.* *Nature* **471**, 225–229 (2011).
4. Brennand, K. J. *et al.* *Nature* **473**, 221–225 (2011).
5. Zhao, T., Zhang, Z.-N., Rong, Z. & Xu, Y. *Nature* advance online publication doi:10.1038/nature10135 (2011).
6. González, F., Boué, S. & Izpisua Belmonte, J. C. *Nature Rev. Genet.* **12**, 231–242 (2011).
7. Kim, D. *et al.* *Cell Stem Cell* **4**, 472–476 (2009).
8. Anokye-Danso, F. *et al.* *Cell Stem Cell* **8**, 376–388 (2011).
9. Kim, K. *et al.* *Nature* **467**, 285–290 (2010).
10. Polo, J. M. *et al.* *Nature Biotechnol.* **28**, 848–855 (2010).
11. Gore, A. *et al.* *Nature* **471**, 63–67 (2011).
12. Laurent, L. C. *et al.* *Cell Stem Cell* **8**, 106–118 (2011).
13. Hussein, S. M. *et al.* *Nature* **471**, 58–62 (2011).
14. Mayshar, Y. *et al.* *Cell Stem Cell* **7**, 521–531 (2010).
15. Lister, R. *et al.* *Nature* **471**, 68–73 (2011).
16. Bock, C. *et al.* *Cell* **144**, 439–452 (2011).
17. Huang, P. *et al.* *Nature* advance online publication doi:10.1038/nature10116 (2011).
18. Marchetto, M. C. N. *et al.* *Cell* **143**, 527–539 (2010).
19. Nguyen, H. N. *et al.* *Cell Stem Cell* **8**, 267–280 (2011).



COMMENT

BUSINESS Roche, Toyota and Nokia on research in the recession **p.277**



THEATRE Echos of Chernobyl in staging of Soviet space race **p.282**

POLICY PhD 'glut' needed to fix problems with food, energy, health and meteors **p.284**

OBITUARY William Nunn Lipscomb Jnr, discoverer of a new type of bonding **p.286**



Atomic attacks would cause huge city fires, like this one in San Francisco in 1906, and smoke would cool the planet.

Nuclear winter is a real and present danger

Models show that even a 'small' nuclear war would cause catastrophic climate change. Such findings must inform policy, says **Alan Robock**.

In the 1980s, discussion and debate about the possibility of a 'nuclear winter' helped to end the arms race between the United States and the Soviet Union. As former Soviet president Mikhail Gorbachev said in an interview in 2000: "Models made by Russian and American scientists showed that a nuclear war would result in a nuclear winter that would be extremely destructive to all life on Earth; the knowledge of that was a great stimulus to us, to people of honour and morality, to act."

As a result, the number of nuclear weapons in the world started to fall, from a peak of about 70,000 in the 1980s to a total of about

22,000 today. In another five years that number could go as low as 5,000, thanks to the New Strategic Arms Reduction Treaty (New START) between the United States and Russia, signed on 8 April 2010.

Yet the environmental threat of nuclear war has not gone away. The world faces the prospect of a smaller, but still catastrophic, nuclear conflict. There are now nine nuclear-weapons states. Use of a fraction of the global nuclear arsenal by anyone, from the superpowers to India versus Pakistan, still presents the largest potential environmental danger to the planet by humans.

That threat is being ignored. One reason for

this denial is that the prospect of a nuclear war is so horrific on so many levels that most people simply look away. Two further reasons are myths that persist among the general public: that the nuclear winter theory has been disproved, and that nuclear winter is no longer a threat. These myths need to be debunked.

The term 'nuclear winter', coined by Carl Sagan and his colleagues in a 1983 paper¹ in *Science*, describes the dramatic effects on the climate caused by smoke from fires ignited by nuclear attacks on cities and industrial areas. In the 1980s my colleagues and I calculated, using the best climate models available at the time, that if one-third of ►

BETTMANN/CORBIS

► the existing arsenal was used, there would be so much smoke that surface temperatures would plummet below freezing around the world for months, killing virtually all plants and producing worldwide famine. More people could die in China from starvation than in the nations actively bombing each other. As many countries around the world realized that a superpower nuclear war would be a disaster for them, they pressured the superpowers to end their arms race. Sagan did a good job of summarizing the policy impacts² in 1984: although weapons were continuing to be built, it would be suicide to use them.

The idea of climatic catastrophe was fought against by those who wanted to keep the nuclear-weapon industry alive, or who supported the growth of nuclear arsenals politically³. Scientifically, there was no real debate about the concept, only about the details. In 1986, atmospheric researchers Stanley Thompson and Stephen Schneider wrote a piece in *Foreign Affairs* appraising the theory⁴ and highlighting what they saw as the patchiness of the effect. They coined the term 'nuclear autumn', noting that it wouldn't be 'winter' everywhere in the aftermath of a nuclear attack. They didn't mean for people to think that it would be all raking leaves and football games, but many members of the public, and some pro-nuclear advocates, preferred to take it that way. The fight over the details of the modelling caused a rift between Sagan and Schneider that never healed. When I bring up the topic of

nuclear winter, people invariably tell me that they think the theory has been disproved.

But research continues to support the original concept. By 2007, models had begun to approximate a realistic atmosphere up to 80 kilometres above Earth's surface, including the stratosphere and mesosphere. This enabled me, and my coauthors, to calculate for the first time that smoke particles would be heated by the Sun and lifted into the upper stratosphere, where they would stay for many years^{5,6}. So the cooling would last for much longer than we originally thought.

DARK DAYS

Many of those who do accept the nuclear-winter concept think that the scenario applies only to a mass conflict, on a scale no longer conceivable in the modern world. This is also false. A 'small' nuclear war between India and Pakistan, with each using 50 Hiroshima-size bombs (far less than 1% of the current arsenal), if dropped on megacity targets in each country would produce climate change unprecedented in recorded human history⁷. Five million tonnes of black carbon smoke would be emitted into the upper troposphere from the burning cities, and then be lofted into the stratosphere by the heat of the Sun. Temperatures would be lower than during the 'Little Ice Age' (1400–1850), during which famine killed millions. For several years, growing seasons would be shortened by weeks in the mid-latitudes (see 'A decade of cooling').

Brian Toon at the University of Colorado in Boulder, Richard Turco at the University of California, Los Angeles, Georgiy Stenchikov at Rutgers University in New Brunswick, New Jersey, and I, all of whom were pioneers in nuclear-winter research in the 1980s, have tried, along with our students, to publicize our results. We have published refereed journal articles, popular pieces in *Physics Today* and *Scientific American*, a policy forum in *Science*, and now this article. But *Foreign Affairs* and *Foreign Policy*, perhaps the two most prominent foreign-policy magazines in English,

"Fidel Castro summoned me to a conference on nuclear winter in Havana."

would not even review articles we submitted. We have had no luck getting attention from the US government. Toon and I visited the US Congress and gave briefings to congressional staff on the sub-

ject two years ago, but nothing happened as a result. The US President's science adviser John Holdren has not responded to our requests — in 2009 and more recently — for consideration of new scientific results in US nuclear policy.

The only interest at a national level I have had was somewhat surreal: in September 2010, Fidel Castro summoned me to a conference on nuclear winter in Havana, to help promote his new view that a nuclear conflict would bring about Armageddon. The next day, my talk — the entire 90 minutes including questions — was broadcast on nationwide television in prime time, and appeared on the front page of the two national newspapers in Cuba.

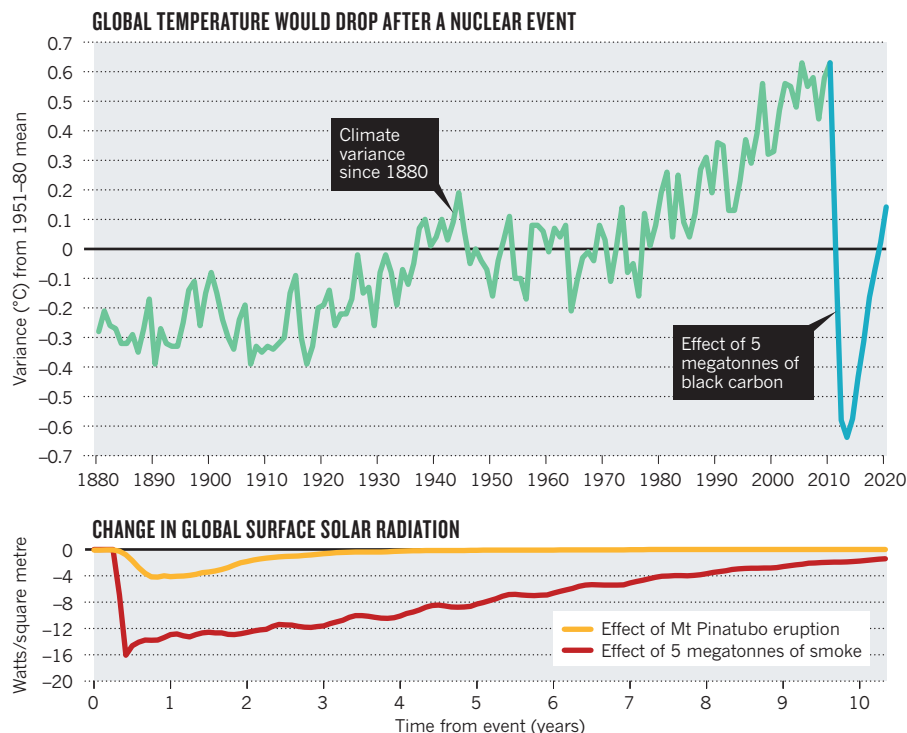
As in the 1980s, it is still too difficult for most people to fully grasp the consequences of a nuclear conflict. But it must be grasped. We scientists must continue to push our results out to the public and to policymakers, so they can in turn push political will in the direction of disarmament. Just as Gorbachev, armed with the knowledge of nuclear winter, helped to end the cold war, so too can the politicians of today use science to support further reductions in arms. The New START treaty is not enough. ■

Alan Robock is in the Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey 08901, USA. e-mail: roboc@envsci.rutgers.edu

1. Turco, R. P., Toon, O. B., Ackerman, T. P., Pollack, J. B. & Sagan, C. *Science* **222**, 1283–1292 (1983).
2. Sagan, C. *Foreign Affairs* **62**, 257–292 (1984).
3. Badash, L. *A Nuclear Winter's Tale: Science and Politics in the 1980s* (MIT Press, 2009).
4. Thompson, S. L. & Schneider, S. H. *Foreign Affairs* **64**, 981–1005 (1986).
5. Robock, A. *et al. Atm. Chem. Phys.* **7**, 2003–2012 (2007).
6. Robock, A., Oman, L. & Stenchikov, G. L. *J. Geophys. Res.* **112**, D13107 (2007).

A DECADE OF COOLING

The detonation of 100 nuclear bombs could cause fires releasing 5 million tonnes of black carbon, with long-term temperature effects — much greater than those from the 1991 eruption of Mount Pinatubo.



SOURCE: REF 5



Medical chemists at pharmaceutical giant Roche want to build better drug-disease models.

Research beyond the recession

Nokia, Toyota and Roche explain how they are weathering the financial crash, the technologies they are investing in and their innovation strategies.

NOKIA

Hard times are opportunities

Leo Kärkkäinen, distinguished scientist, Nokia Research Center, Helsinki, Finland

Different parts of the research and development (R&D) chain respond to hard times in different ways. Agility, out of the box thinking and solid R&D performance are even more highly valued now than in times of plenty. However, in a deep recession, no one is insulated from fiscal pressures. The first parts of the R&D chain, open innovation and long-term research, are asked to 'fail fast' — that is, to figure out the future relevance of new technologies as soon as possible. For

the development part, it is necessary to have solid product delivery and agility in responding to fast-moving consumer trends to steer the product towards markets that are not as sensitive to recession pressures.

Yet hard times are often the best times to invest and create new products, because they are precisely when the competition is less likely to be able to respond to the challenge. So cutting R&D costs under fiscal pressure is sometimes, paradoxically, the most expensive thing to do.

In the next 20 years, we will see the rise of carbon-based electronics, using ultrathin graphene sheets, which will change how electronic components are produced and integrated onto chips or devices. Graphene is thin and transparent, often created by peeling away an atomic layer of carbon from a growing substrate like copper.

NATURE.COM
Pfizer shuts UK
research site:
go.nature.com/x2ghpn

This will make it naturally flexible, enabling a new breed of electronic devices, embedded in everyday objects, and wearable like cloth. These flexible electronic devices will be able to talk to each other to provide a better user experience, from faster response times to context-sensitive behaviour.

At the same time, manufacturing methods will change dramatically, driving down the cost of the products — we will see more roll-to-roll printed electronics, as already used for flexible displays. This could increase technology adoption rates in poorer nations, affecting global social development. By 2020 we can also expect widespread applications of artificial intelligence, self-driving cars and robots that interact with humans more easily.

Nokia is driven by a passion for doing things, and a positive company culture is very important to us. We have companywide activities to encourage diversity in innovation. In our yearly Nokia Excellence Award, hundreds of well performing and successful innovative projects are reviewed and the best dozen are personally evaluated by the Nokia chief executive and board members in a face-to-face event with the inventors.

One of the biggest challenges in innovation is prioritization: new ideas may not have a convincing enough business case or value until changes in the business environment make them obvious to everyone (including competitors). This can, of course, happen when it is too late to reap any rewards. One way we tackle this is to combine prioritization with crowd-sourcing of new ideas from our employees. In our Nokia Sphere project, employees can vote to work and improve on ideas that seem promising, driving implementation from the bottom up.

TOYOTA

Improve efficiency of development

Toyota Motor Corporation, Aichi, Japan

We have been working in a very difficult business environment since the recession began in 2009 (see 'Corporate changes'). But we are still investing more than 700 billion yen (US\$8.7 billion) a year in R&D. We want to keep our competitive edge in technology and products, so we are maintaining high levels of investment to develop advanced technology related to safety, the environment and energy. We've used the recession as an opportunity to work with our suppliers to improve our development efficiency and we hope to take advantage of this in the future.

At the moment, there is no clear ▶

► alternative to petrol as an energy source for cars. So we are developing a wide range of products based on hybrid-vehicle technology, combining an electric motor and a petrol engine. Our approach is to develop the best cars for the consumer in each different market.

We currently have a strong focus on batteries for future electric vehicles. Although lithium-ion batteries are becoming more widely used, it is hard to see electric vehicles completely replacing conventional passenger cars, even if we push the performance of lithium-ion batteries to the limits. We have to solve problems of energy storage density and cost. We are researching and developing all-solid and metal-air batteries, which are two promising alternatives to lithium-ion.

Another possible game-changing technology is solar power. More and more households are using solar cells. At the moment, some of our hybrid Prius cars have solar-powered ventilation systems that operate while the car is parked, but it may also be possible to use solar power to drive the vehicle if we can achieve a breakthrough in the efficiency of generating electricity from solar energy.

In the long term, we believe that the use of vehicle telematics will revolutionize the car industry. We are seeing rapid development and innovation in automated driving and accident prevention. As vehicle-control technology advances, more cars may be able to avoid collisions. Then it may become possible to change vehicle structures and make cars much lighter. That will in itself reduce energy usage.

The Japanese idea of *monozukuri*, which could be translated as making things, is at

the heart of Toyota's approach. We think that new ideas are created by digging into the root causes of problems and by finding out facts through *genchi genbutsu*, which means actually going to a site and discovering the real situation for yourself. It is important that we nurture our employees to take this practice to heart. For the past 50 years, this approach has been the driving force behind the innovation and originality in our development processes.

ROCHE Collaborate with the public sector

Jean Jacques Garaud, global head of pharma research and early development, Roche Holding, Basel, Switzerland

The recession is diminishing the funding available for research at publicly funded scientific institutions. This compels them to be more open to, and more collaborative in, public-private partnerships. Since the integration into Roche of Genentech, a Californian biotechnology company, in 2009, Roche has operated two autonomous Research and Early Development units, pRED and gRED, with distinctive approaches. In the first 18 months of pRED, we've developed and driven external collaborations, ranging from relationships with individual academics to entire networks with leading academic and health institutions.

At the same time the economic crisis increases the pressure on drug prices and

forces us to home in on drug candidates that will add value from a medical and public-health standpoint. We are focusing efforts on personalized health care, because patients with the same condition can react to the same treatment in different ways — and sometimes even receive treatment that is inappropriate for them. To better fit the treatment to the patient, we must concentrate on better understanding the molecular basis of diseases and their heterogeneity.

I'm optimistic that these recessionary challenges can be turned into opportunities to make health care better, safer and more effective.

Our ultimate goal is to understand the biology of diseases and translate this knowledge into the clinic. New technologies that will help include cell-penetrating peptides that may allow the delivery of drugs into cells as well as therapeutic interactions on the cell surface. For peptides in general, we will need to develop synthesis methods to overcome difficulties, such as structural instability, that can weaken peptide interaction with targets and reduce activity and specificity.

Stem cells will also be increasingly important as translational-research tools. With differentiated cells derived from stem cells, we are able to study the effects of drug compounds on clinically relevant targets and observe cellular functions at an early stage.

Finally, computer modelling and simulation could also be game changers, if we can build more reliable drug-disease models to better design experiments and predict their outcome.

To encourage such innovation, Roche fosters an environment that allows our scientists to grow and experiment with new ideas and approaches. One way to do that is to talk about science itself, not just about managing science. We have launched a 'barn initiative' to provide informal environments for kindling creativity in settings from campuses and castles to converted barns. At these 'barns', away from their day-to-day projects, scientists can engage in positive and challenging scientific discussions on a specific theme.

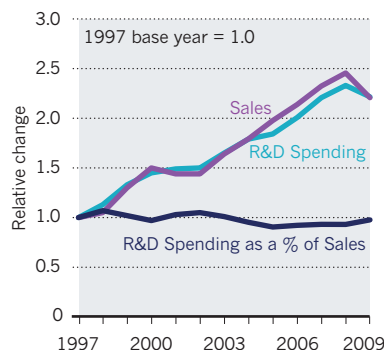
It is also important to provide the recognition and the rewards that scientists deserve. Our publication strategy explicitly encourages publishing in scientific journals and we advocate the exchange of ideas at scientific conferences. ■

CORPORATE CHANGES

In 2009, corporate research and development (R&D) spending declined for the first year in more than a decade (see graph), according to a study of 1,000 of the world's most research-intensive companies by New York analysts Booz & Company.

Total R&D spending in 2009 dropped 3.5%, but revenues fell more sharply, by 11%. So R&D is still one of the last places that corporations make cuts. About half of the 1,000 firms cut their R&D portfolio in 2009, but nearly all the cuts came in three industries: car manufacturers, computing and electronics.

R&D AND SALES



THE INNOVATION TOP 5

2009 rank	2009 R&D spend US\$ millions (2008 rank in parentheses)	Change from 2008	As a % of Sales
1	ROCHE (3) \$9,120 m	▲ 11.6%	
2	MICROSOFT (4) \$9,010 m	▲ 10.4%	
3	NOKIA (2) \$8,240 m	▼ 1.0%	
4	TOYOTA (1) \$7,822 m	▼ 19.8%	
5	PFIZER (6) \$7,739 m	▼ 2.6%	

Research spending in the health-care sector grew by a modest 1.5% in 2009, as reflected in the rankings of the top spenders (see table). Toyota Motor Corporation and Nokia both dropped, while Roche Holding climbed two places to take the top spot ahead of Microsoft.

CORRECTION

In the Comment article 'The art of conservation' (*Nature* **472**, 287–289; 2011), the 1964 Durrell Wildlife Conservation Trust logo and the 1961 Friends of the Earth International logo were actually from 1999 and the 1970s, respectively.

FROM THE 1949 FILM *A RUN FOR YOUR MONEY*. COURTESY RONALD GRANT ARCHIVE

Better understanding of decision-making processes in the brain might predict which perpetrators will offend again.

NEUROSCIENCE

My brain made me do it

Adam Kepecs urges caution in considering the unconscious mind in the justice system.

A surprising view has been gathering momentum in neuroscience: most of our thoughts and actions are driven by unconscious brain processes that are hidden from conscious introspection. So if consciousness is rarely in the driver's seat, and if we cannot choose our genes or the childhood experiences whose interactions form our brains, then are we responsible for our actions?

In *Incognito*, accomplished neuroscientist David Eagleman — author of the best-selling short-story collection *Sum* (Canongate, 2010) — examines this gap between our conscious and unconscious selves. He offers a whirlwind of stories, from visual illusions and sleep-walking killers to ovulating strippers, all carefully chosen to drive home his main point that our brains “neurally preordain” us to make decisions. As is common in books aimed at a general readership, the intriguing and sometimes bizarre case studies create a tension between

journalistic musings and more detailed arguments. Although specialists may feel that the balance tilts toward the journalistic, Eagleman's expertise comes through.

Since Sigmund Freud's famous psychological framing of the unconscious in the late nineteenth century, modern neuro-

science has shown that most processing in the brain is unconscious. We are unaware of routine processes and have little insight into our choices and preferences. For instance, men unknowingly prefer photographs of women with dilated pupils, presumably because male brains evolved algorithms to recognize pupil dilation as an indicator

of sexual arousal. In another experiment, people's descriptions of the strategies they used to make simple economic decisions differed from the rules that they actually used, suggesting that their conscious explanations were formed post hoc and without access to their decision-making process. Through such examples, Eagleman demonstrates that unconscious processes can be clever, adaptive and even outperform the best computer algorithms.

If our brains can carry out such amazing feats without us knowing, why have consciousness at all? Eagleman answers this question with a metaphor. Consciousness, he says, is like the chief executive of a large company. He or she has little knowledge of the day-to-day operations, yet is indispensable for setting goals and arbitrating between conflicting departments. Similarly,



Incognito: The Secret Lives of the Brain

DAVID EAGLEMAN
Pantheon/Canongate:
2011. 304 pp./272 pp.
\$26.95/£20

NATURE.COM
See *Nature's* special
issue on Science in
Court:
go.nature.com/ez6pwk

consciousness gets only the abridged, delayed and sometimes contradictory reports from neural subroutines. And, much like a chief executive trying to explain him- or herself to the board of directors, consciousness will “fabricate stories to explain the sometimes inexplicable dynamics of subsystems in the brain”.

Having described the hidden life of our brain circuits, Eagleman moves to an original and provocative discussion of the legal consequences of the unconscious decider within us. Imagine two defendants on trial for murder: one has a large brain tumour next to an area associated with aggression, whereas the other one shows no obvious change in his brain. Most people would not hold the first defendant responsible for his actions. Eagleman argues that as we gain a better understanding of the biology of decision-making, we will be forced to conclude that all crime is caused by faulty brain circuits arising from genetic and environmental interactions over which the perpetrator has no control.

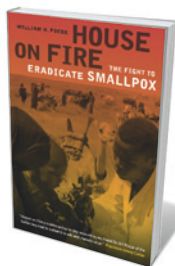
An improved understanding of how subtle changes in the brain generate deviant behaviour would therefore extend the insanity defence — ‘my brain made me do it’. Eagleman suggests that a forward-looking legal system should consider biological information to predict how likely a person is to commit a crime again, and take this into account for sentencing. As most criminals commit offences because they are unable to inhibit their impulses, Eagleman proposes that rehabilitative “prefrontal workouts”, aimed at improving self-control, should be a mainstay of the justice system. Crime would still land you in jail, but the focus would be on protecting society, not on punishment.

My feeling is that we need to be extremely cautious in advancing such a brain-centric legal system. A world in which judges are instructed to consider the genetics and neural make-up of defendants, as Eagleman advocates, evokes Phillip K. Dick’s short story *The Minority Report*. If sentencing decisions consider the biological likelihood of recommitting a crime, it is easy to imagine the next step of considering preventive measures before a crime has been committed — a kind of ‘Department of Precrime’.

Whether or not one agrees with Eagleman, discussions about these difficult issues at the intersection of neuroscience and society are essential and timely. He should be lauded for his clear exposition of the consequences of our emerging understanding of the brain. *Incognito* is a smart, captivating book that will give you a prefrontal workout. ■

Adam Kepecs is assistant professor of neuroscience at Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA.
e-mail: kepecs@cshl.edu

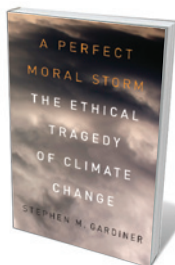
Books in brief



House on Fire: The Fight to Eradicate Smallpox

William H. Foege UNIVERSITY OF CALIFORNIA PRESS 240 pp. \$29.95 (2011)

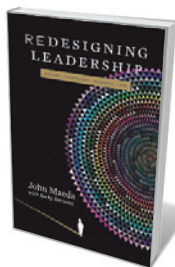
Adding to the series of California/Milbank Books on Health and the Public, this part-memoir, part-history by epidemiologist William Foege recounts his involvement in the global vaccination programmes that eradicated smallpox in the 1960s and 1970s. Foege, now a senior fellow at the Bill & Melinda Gates Foundation in Seattle, Washington, reflects on the strategies that led to wide uptake of the vaccines across Africa and India and discusses their successes and vulnerabilities.



A Perfect Moral Storm: The Ethical Tragedy of Climate Change

Stephen M. Gardiner OXFORD UNIVERSITY PRESS 512 pp. £22.50 (2011)

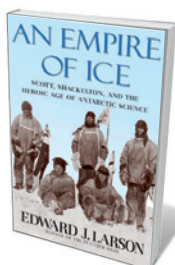
Inaction on climate change is more than a political or explanatory bungle — it is a moral failure, declares philosopher Stephen Gardiner. He identifies three reasons for this: the global imbalance of power, such that rich nations dominate poor ones; intergenerational factors, such that present generations dictate the world that future ones will inhabit; and our inability to make predictions using current scientific knowledge. A ‘perfect storm’ of these three factors creates an ethical headache, Gardiner contends.



Redesigning Leadership: Design, Technology, Business, Life

John Maeda with Becky Bermont MIT PRESS 104 pp. \$18 (2011)

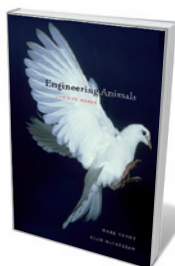
Celebrated designer and computer scientist John Maeda shares his thoughts on leadership in this concise volume. Interspersing his musings with philosophical tweets, he reflects on how he has sought out imaginative ways to run teams and organizations, from the Media Lab at the Massachusetts Institute of Technology in Cambridge to the Rhode Island School of Design in Providence. He describes how to make meetings run faster and be more fun, the team-building benefits of free food and how to harness conflicting opinions within a group of creatives.



An Empire of Ice: Scott, Shackleton, and the Heroic Age of Antarctic Science

Edward J. Larson YALE UNIVERSITY PRESS 360 pp. £18.99 (2011)

Looking at the broader context of the race to be first to reach the South Pole in the early twentieth century, historian Edward Larson celebrates the centenary of these explorers’ achievements. It was the greater scientific ambition of the British Antarctic expeditions, he argues, that caused Robert Scott and Ernest Shackleton to be slower than their more narrowly focused Norwegian rival, Roald Amundsen, who made it to the South Pole 35 days before Scott and his ill-fated team.



Engineering Animals: How Life Works

Mark Denny and Alan McFadzean BELKNAP PRESS/HARVARD UNIVERSITY PRESS 400 pp. \$35 (2011)

From soaring albatrosses to croaking bullfrogs, different creatures exploit various aspects of engineering to help them fly, hunt or communicate. In a clear and well-illustrated account, former aerospace engineers Mark Denny and Alan McFadzean describe the principles of physics that underlie animals’ sense of smell, their use of sonar, and how they flock, signal to each other and consume energy.



Sergei Korolyov (played by Darrell D'Silva) was celebrated after the success of the Sputnik satellite.

THEATRE

The chief designer

A clever play shows how engineer Sergei Korolyov drove the Soviet space programme, finds **Philip Ball**.

It is a curious year of anniversaries for the former Soviet military-industrial complex. Fifty years ago, the cosmonaut Yuri Gagarin became the first person in space, orbiting the world for 108 minutes in the *Vostok 1* spacecraft. And 25 years ago, reactor four at the Chernobyl nuclear plant in Ukraine exploded, sending a cloud of radioactive debris across northern Europe.

One triumph, one failure. Rona Munro's play *Little Eagles*, commissioned by the Royal Shakespeare Company for the Gagarin anniversary, understandably makes no mention of the later disaster, but connections assert themselves throughout.

Both events were fruits of the cold war nuclear age. The rockets made by Sergei Korolyov — chief architect of the Soviet space programme and the play's central character — armed President Nikita Khrushchev with intercontinental ballistic missiles before they took Gagarin to the stars. But the space programme degenerated along the same bureaucratic lines that made an exclusion zone of Chernobyl.

Impossible demands from technically clueless party officials, and

the terror of engineers and operatives at the consequences of neglecting them, eventually compromised both technologies fatally — most notably in *Little Eagles* in the crash of *Soyuz 1* in 1967, which killed cosmonaut Vladimir Komarov. Gagarin was listed as the backup pilot for that mission, but was by then too valuable a trophy to be risked in another space flight. All the same, Gagarin died a year later, during the routine training flight of a jet fighter.

Callous disregard for life marks Munro's play throughout. We first see Korolyov, a military rocket engineer, in the Siberian labour camp where he was sent during Stalin's purge of the officer class just before the Second World War. As the Soviets developed their rocket programme, the stupidity of sending someone so brilliant to a virtual death sentence dawned on the regime, and Korolyov was freed. During the 1950s, he gained control of the whole space enterprise, becoming known as the Chief Designer.

Munro's portrayal of Korolyov seems accurate, if the testimony of one of his chief scientists is anything to go by: "He was a king,

Little Eagles

WRITTEN BY RONA MUNRO. DIRECTED BY ROXANA SILBERT. Hampstead Theatre, London. 16 April–7 May 2011.

a strong-willed purposeful person who knew exactly what he wanted ... he swore at you, but he never insulted you. The truth is, everybody loved him." Portrayed in the Hampstead Theatre production by Darrell D'Silva, he is a swaggering, cunning, charming force of nature, playing the system so that he might realize his dream of reaching for the stars. He reciprocates the love of his 'little eagles', the cosmonauts chosen with an eye on the *Vostok* capsule's height restrictions. (Even today the *Soyuz* capsule excludes tall cosmonauts.)

But for the leaders of the Soviet Union, rocketry was merely weaponry, a way of demonstrating superiority over their foes in the West. Korolyov becomes a hero for beating the Americans with Sputnik, and then for *Vostok 1*. But when the thuggish, foul-mouthed Khrushchev (played by a terrifying Brian Doherty) is retired in 1964 in favour of Leonid Brezhnev, the game changes. Brezhnev sees no virtue in Korolyov's dream of a Mars mission, and worries instead that the Americans will beat them to the Moon. The rushed and bungled *Soyuz 1*, launched after Korolyov's death in 1966, was the result.

Out of this fascinating but chewy material, Munro has woven a moving and often beautiful tale. Gagarin's own story is a subplot. Grounded as a toy of the Politburo, we see his sad descent into the vodka bottle but not his ignominious end — that is too much to be shoehorned into this packed play. Nevertheless, it is a satisfying and wise production.

The play might be regarded as a foil to *The Right Stuff*, Tom Wolfe's 1979 account of the US space programme, which was made into an exhilarating film in 1983. Wolfe's celebration was a fitting tribute to the courage and ingenuity that took humans to the Moon, but a glimpse at the other side of the coin was long overdue. There is something awesome as well as awful in the grinding resolve of the Soviets to win the space race relying on just the chief engineer, "convicts and some university students", as Munro puts it.

Little Eagles shows us the mix of noble and ignoble impulses in the space race that the US programme, with its Columbus rhetoric, still cannot afford to acknowledge. The play recognizes the glory of seeing the stars and Earth from beyond the atmosphere, but reveals the human space-flight programmes as a product of their tense times, as nationalistic black holes for dollars and roubles (and now yuan too). And, like Chernobyl, such politically motivated displays are open to mistakes not from excessive ambition but from fear of failure.

Crucially, Munro leaves the final judgement to the audience. "They say you changed the whole sky and everything under it," Korolyov's physician says to him in the final lines. "What does that mean?" ■

Philip Ball is a writer based in London.

NATURE.COM

Nature's special issue on Russian science:

go.nature.com/cwx8g



Cosmology is a strong theme within the garden of landscape designer Charles Jencks.

Q&A Charles Jencks

Cosmic gardener

Charles Jencks designs landscapes and sculptures to convey concepts in astronomy, biology and mathematics — notably at CERN, Europe's particle-physics lab near Geneva, Switzerland, and in his Garden of Cosmic Speculation near Dumfries in Scotland, UK. On the launch of his new book, he discusses green architecture and metaphor.

Why use landforms and landscapes to express scientific ideas?

My book *The Universe in the Landscape* describes my designs at all scales, from small gardens to a restored open-cast coal mine. They mix different media — architecture, sculpture, planting and epigraphy — to interpret basic ideas of the cosmos. I see turf mounds as a medium through which we can interpret a larger cosmic nature. This endeavour parallels those in prehistory, when people made landforms such as the stone circles of Brodgar in the Orkney Islands and Stonehenge near Salisbury, UK.

How did you get involved with CERN?

CERN's director-general, Rolf Heuer, and his team asked me to collaborate on a project at the centre of the Large Hadron Collider. They had built a wooden dome, which they call the Globe of Science and Innovation, and I argued that they should make a

The Universe in the Landscape: Landforms by Charles Jencks
CHARLES JENCKS
Frances Lincoln: 2011.
256 pp. £40

landscape around it. I talked to many scientists, including the astronomers Martin Rees and Bernard Carr, who suggested using uroboros, the

What will the CERN landscape look like?

As an overall strategy, a connected landscape will protect what little green land remains around the Globe. I have proposed a design in the shape of a modified uroboros — a ring that connects steps in the relative scale of objects in the Universe, with a question mark for the snake's head. CERN's particle collisions have become a sculptural icon used for

architectural details: in the form of an eye, with rays exploding outwards from its central point. The construction could be finished in three years.

How does your Garden of Cosmic Speculation explore cosmology?

There is a 'Universe cascade' where water runs down and time runs up, in a series of stone steps that show the history of the Universe. As you walk up the steps, jumping from platform to platform, you see the Universe slowly unfolding from its origins in the mists of the quantum soup. Mathematical physicist Roger Penrose asked me, "How could you build superstrings in concrete when they are so clearly going to be proven wrong in ten years?" But you enjoy the play of uncertainty in a garden. You can revise your work when you make mistakes. Besides, they date the design precisely at the point we believed these theories.

What was your intention in building it?

The garden is a project started in 1988 with my late wife, Maggie. We tried to translate some of the metaphors of science into landscape design. There is a DNA garden with six cells whose walls are made up of various plants. At the centre of each cell is the nucleus, and at the centre of the nucleus are six versions of the double helix, unfolding with RNA coming out to be read by plants that represent ribosomes. It is a critical look at the relationship between DNA and the cell.

How is science changing architecture?

Recent attempts to pull together complexity theory and architecture have led to an enormous amount of design, but not much of it has been built or is convincing. However, computers are transforming architecture. Architect Frank Gehry uses software written by French aerospace engineers to design curved buildings. And Zaha Hadid and her students in London are putting forward 'parametric architecture', which uses rules to generate appropriate designs.

How do you see the role of architects in future?

Every time an architect designs a building, they are predicting what will be relevant for the world in the coming decades. They must also persuade clients and society. Foster and Partners, one of the greenest architecture firms on the planet, got the Reichstag in Berlin running on vegetable oil. But they also build expensive skyscrapers. The problem for the profession is that architects do not control enough of the building process to lead the green agenda. But they should try. You fight the right battles even if you do not expect to win the war. ■

INTERVIEW BY JASCHA HOFFMAN

CORRESPONDENCE

Guidelines for HIV in court cases

In many nations it is a crime to infect someone with HIV by intention or non-disclosure. As phylogenetic experts who advise courts worldwide, we are calling for guidelines on how phylogenetics should be used in criminal HIV investigations. The inappropriate use of such evidence in suspected transmission cases can have dire legal and social ramifications.

The scientist's job is not to argue for or against a defendant's guilt: that is a task for lawyers. Phylogenetic investigators should limit themselves to an expert opinion on what information about viral transmission can be deduced from their analysis. This must be derived impartially, for example by blinding the identities of case subjects.

Scientists must explain to courts that phylogenetic analysis cannot 'prove' any particular hypothesis, such as 'person A infected person B'. Rather, results may be compatible with several hypotheses, or support one over another.

An *a priori* hypothesis should be formulated by different independent epidemiological experts, based on contact possibilities between the purported victim(s) and the defendant, and on any additional contacts or risk factors.

Phylogenetic analysis alone cannot exclude the possibility that HIV was transmitted from A to B through unsampled persons. Although the direction of viral transmission can sometimes be supported, it does not prove direct transmission.

Thomas Leitner on behalf of 8 co-authors*, Los Alamos National Laboratory, New Mexico, USA.
tkl@lanl.gov

*A full list of co-authors is available online at <http://dx.doi.org/10.1038/473284a>.



PhDs: what's left if science abdicates?

I disagree that we have a glut of scientists with PhDs (www.nature.com/phdfuture). The corporate view of PhD numbers in terms of what the market will bear ignores the major problems that only science can solve in the coming century.

The list is long: natural disasters, such as earthquakes and incoming celestial objects; environmental degradation; sustainable energy; famine and violence; untreatable medical conditions; and threats such as antibiotic resistance. If science abdicates, there is nothing else.

The urgency of these problems requires a large cadre of trained individuals to be enlisted to defend our planet. The size of the military is dictated by our defence needs, not the market. In science, by analogy, our global defence needs are soaring.

Spending a few years in the service of science and the greater good, being rewarded with an advanced degree and, for example, going on to teach in high schools is an honourable fix.

Kenneth S. Kosik University of California, Santa Barbara, USA.
kosik@lifesci.ucsb.edu

➔ **NATURE.COM**

Join the debate on the future of the PhD:
go.nature.com/phdfuture

PhDs: Israel also trains plenty

You contend that few PhDs are trained in the Middle East outside Egypt (*Nature* **472**, 276–279; 2011). Israel is a sizeable contributor as well.

In 2008–09, Israel had more than 10,000 students enrolled in doctoral programmes (Central Bureau of Statistics, Israel). This is fewer than Egypt's 35,000 for the same period, but many more per capita.

Given ongoing tensions in the region, the scientific press has a responsibility to report data related to higher education and research transparently and accurately (see also *Nature* **471**, 37; 2011).

Thomas Hays Mount Sinai School of Medicine, New York, USA.
thomas.hays@mssm.edu

Crop failure signals biodiversity crisis

Crop failures have pushed up food prices globally (*Nature* **472**, 169; 2011). Human well-being depends on biodiversity and natural habitats as a source of food. Ironically, the countries harbouring these vital natural assets are also those currently facing the most severe food crises.

A report from the investment bank Nomura (go.nature.com/pwrlc9) introduces a global index for measuring nations' food vulnerability. The most vulnerable depend totally on imported food, and citizens spend more than one-third of their salaries on it.

Of the 35 most vulnerable countries, 15 contain tropical biodiversity hotspots. To produce more food, these countries may lease out their biodiversity-rich land to farm cash crops. Liberia, for example, intends to add 220,000 hectares of oil-palm

plantation (go.nature.com/xblcjz) to its existing 1.6 million hectares of agricultural land in the southeast, one of the last strongholds of tropical forest in western Africa.

Vulnerable nations need better cooperation among governments to address the structural causes of imbalances in the international agricultural system; more research into new technologies that incorporate the food-production requirements of the rural poor; and stronger protection of natural systems by linking biodiversity preservation to increased food security.

Kelvin S.-H. Peh University of Cambridge, UK.
kelvin.peh@gmail.com

China must reduce fertilizer use too

Environmental damage caused by reactive nitrogen is not just a European problem (*Nature* **472**, 159–161; 2011). China must also rein in its overuse of nitrogen fertilizers — which accounts for 40% of global production since 2006 — to balance food-security requirements with the protection of human health and the environment.

Despite China's nitrogen consumption almost doubling between 1990 and 2009, its grain production increased by just 22%. Although the research community widely recognizes the problem of fertilizer overuse, farmers in China continue the practice, which is promoted by some agricultural-extension advisers and by sellers of fertilizer.

Chinese farmers need to be taught how, when and in what quantities fertilizer should be applied. The existing agricultural-extension system must revert to its role of assisting farmers by methods other than promoting fertilizer sales.

Establishing an environmental-extension system at the township level could also help to prevent overuse of fertilizers and pesticides.

Peng Gong, Lu Liang and Qiang Zhang *Institute for Global Change Studies, Tsinghua University, China.*

penggong@tsinghua.edu.cn

Data archiving is a good investment

We have found that ongoing financial investment in data-archiving infrastructure yields an impressive scientific return, and believe that it should be wholeheartedly supported by research funding agencies (see, for example, go.nature.com/nzftf3).

We used Dryad (see <http://datadryad.org>), an international, open, cost-effective data repository for the biological sciences, to estimate the cost of archiving data from more than 10,000 publications. We found that these could be curated and the data preserved at an annual cost of about US\$400,000.

As an example of how much research is typically published per grant dollar, core grants in population and community ecology from the US National Science Foundation averaged 3–4 publications per \$100,000 of grant between 2000 and 2005 (S. Reyes, A. Tessier and S. Mazer, unpublished results). That is, \$400,000 invested in original research resulted in about 16 papers.

Dryad cannot yet tell us how effective data archives are in facilitating primary research publications, but the Gene Expression Omnibus (GEO) database at the US National Center for Biotechnology Information offers some insight. To estimate data reuse, we searched the full text of articles in PubMed Central for mention of any of the 2,711 data sets deposited in GEO in 2007. We excluded articles whose authors' names overlapped with those depositing the data set. Extrapolating the 338 hits in PubMed Central to all of PubMed, we estimate that the

GEO 2007 data sets made third-party contributions to more than 1,150 published articles by the end of 2010, and reuse continues to accumulate rapidly (H. A. Piwowar, T. J. Vision and M. C. Whitlock *Dryad Digital Repository* doi:10.5061/dryad.j1fd7; 2011).

Assuming that Dryad has a comparable rate of reuse and collects at least 2,500 data sets annually, an investment of \$400,000 in one year should contribute to more than 1,000 papers in the next four years — far more than the accepted value for a research dollar.

Heather A. Piwowar *Dryad, and the National Evolutionary Synthesis Center, Durham, North Carolina, USA.*

hpwowar@nescent.org

Todd J. Vision *Dryad, and the University of North Carolina, Chapel Hill, North Carolina, USA.*

Michael C. Whitlock *Dryad, and the University of British Columbia, Vancouver, Canada.*

Noisy oil exploration disrupts marine life

Fossil-fuel operations in the Arctic will inevitably compromise habitat — regardless of spills (*Nature* **472**, 163; 2011).

The seismic airgun surveys used for hydrocarbon exploration and for monitoring deposit conditions can disrupt foraging behaviour of bowhead whales at long distances. They also seriously diminish fisheries catches of haddock and other Arctic species, and have halted the migration of fin whales (a non-Arctic species) at a range of more than 175 kilometres.

Bowhead and beluga whales avoid oil-derrick operations. Many other noises associated with fossil-fuel exploration and production — such as construction, shipping, transport helicopters and underwater acoustic telemetry — have a deleterious impact on the marine acoustic environment.

We do not yet know about the impacts of noise from thruster-stabilized exploration

platforms and from sea-floor processing equipment such as wellhead chokes, separators and re-injectors, which operate out of sight and under extreme pressures.

In collaboration with the World Wildlife Fund and the Natural Resources Defense Council, we are developing a peer-reviewed website that can be understood by a lay audience in order to explore some of these issues (see go.nature.com/5vuebe).

Michael Stocker *Ocean Conservation Research, Lagunitas, California, USA.*
mstocker@msa-design.com

Address education inequality in India

Narrowing the educational achievement gap between different social groups in India remains a major challenge, despite 60 years of affirmative-action policy (*Nature* **472**, 24–26; 2011). Using publicly available data from the country's top medical school, the All India Institute of Medical Sciences (AIIMS), we found that performance was poor among students admitted under a government scheme for socially disadvantaged groups.

All government and government-aided institutions in India allocate a fixed percentage of places on educational courses to socially and economically disadvantaged students. But in 1995–2005, out of more than 600 indigenous tribes with access to such positions, one small group from northern India accounted for 36% of students admitted to the AIIMS.

Between 1998 and 2006, socially deprived students accepted into the AIIMS scored 13.6% less in the entrance exam than students from non-disadvantaged social classes ($P < 0.001$). In 1989–98, such students also had double the dropout rate of non-disadvantaged students (6% versus 3%; $P > 0.05$). In the ten years for which data are available (1995–2005), 61.4% of students admitted to

government-reserved positions had to resit examinations in at least one subject, compared with 15.2% of non-disadvantaged students ($P < 0.001$).

To address such inequality, India should adopt measures that have proved successful in other countries. These include wider access to quality primary education; standardized assessment of students; and academic support for students who are lagging behind. More research to assess this inequality is also needed to inform education policy.

Manas Kaushik, Subha Ramani *Boston University, USA.*
mkaushik@post.harvard.edu

Fund experiments on atmospheric hazards

The radioactivity released from Japan's damaged Fukushima Daiichi nuclear power plant has increased the urgency to fund tracer experiments that will improve models of atmospheric dispersion and reinforce confidence in emergency procedures.

The last major tracer experiments were conducted in the mid-1990s. So the predictive capabilities of current atmospheric-dispersion models have not been properly tested, hindering their evaluation and development.

To generate more observational data, multiple-scale atmospheric tracer experiments should use non-hazardous, climate-neutral substances and a realistic release term with varying source strengths. Modellers could estimate emissions in real time using a limited set of observations without knowing the actual release rates, and later improve their models and data-reconstruction methods on the basis of the real source terms and measurements.

Stefano Galmarini *European Commission Joint Research Centre, Italy.* **Andreas Stohl** *Norwegian Institute for Air Research, Norway.* **Gerhard Wotawa** *Central Institute for Meteorology and Geodynamics, Austria.*
gerhard.wotawa@zamg.ac.at

William Nunn Lipscomb Jr

(1919–2011)

Chemist who discovered a new kind of bonding.

William Nunn Lipscomb Jr could have made a career in music or science. He was an accomplished clarinetist who played Mozart with ease and grace, and attended the University of Kentucky in Lexington on a music scholarship. It is chemistry's good fortune that he ultimately chose science. His work on the boron hydrides led to a major rethink of how atoms bind together to form stable molecules.

Lipscomb, who died on 14 April, was born in Cleveland, Ohio, to a physician father and housewife mother. The family moved to Lexington when he was a year old. Both his grandfather and great-grandfather had been physicians, and Lipscomb was expected to continue the family tradition. But after graduating with a degree in chemistry from the University of Kentucky in 1941, he entered the graduate programme in physics at the California Institute of Technology in Pasadena.

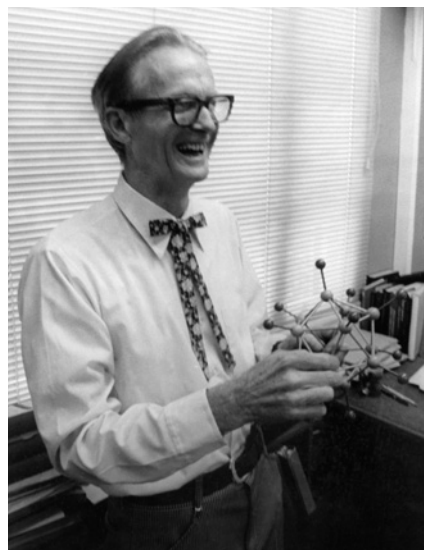
Lipscomb soon returned to chemistry and under the influence of his mentor, Nobel laureate Linus Pauling, he developed an intense interest in chemical bonding. After completing his PhD in structural chemistry (during which he also conducted classified Second World War related research that involved, as he recounted, "walking around with beakers of nitroglycerine"), he joined the faculty at the University of Minnesota in Minneapolis in 1946. In 1959, he was appointed professor of chemistry at Harvard University in Cambridge, Massachusetts.

Lipscomb's research from 1960 onwards included important structural studies of enzymes. But it was his investigations of the boron hydrides, or boranes, from the late 1940s until the 1970s that led to his being the sole recipient of the 1976 Nobel Prize in Chemistry.

The impact of this work is best appreciated in light of the ideas on chemical bonding that prevailed at the time. Organic synthesis, a highly successful branch of chemistry concerned with manipulating hydrocarbons and their derivatives, is based on the perception that carbon atoms bind to other atoms through covalent bonds consisting of a pair of electrons. It was assumed that boron, a neighbour of carbon in the periodic table, would behave similarly, but its hydrides posed a major problem.

Boron's unconventional chemistry had fascinated and confounded researchers for

years. At first, following its isolation in 1808, the element seemed unremarkable, forming the expected trivalent compounds such as BCl_3 and $\text{B}(\text{CH}_3)_3$. Its simplest hydride, it was thought, must be BH_3 . William Ramsay, who had won a Nobel prize in 1904 for his discovery of the noble gases, was convinced of this. Not so the German chemist Alfred Stock, who took up the study of boranes in 1909. Using an ingenious vacuum glass apparatus of his own invention, Stock prepared and isolated a whole family of boron hydrides, none



of which was BH_3 . In fact, he painstakingly proved that the simplest borane was B_2H_6 (diborane). For many years the molecular structures of these hydrides, which included B_4H_{10} , B_5H_9 , B_6H_{10} and $\text{B}_{10}\text{H}_{14}$, among others, remained unknown. In 1948, the basket shape of $\text{B}_{10}\text{H}_{14}$, which is a solid at room temperature, was established by the US chemist John Kasper and his colleagues.

A general understanding of borane structures emerged only when Lipscomb and his co-workers investigated the hydrides using both theory and X-ray crystallography. The challenge was formidable, as the boranes smaller than $\text{B}_{10}\text{H}_{14}$ are gases or volatile liquids at room temperature. This meant that the crystals had to be grown in sealed capillaries at very cold temperatures using liquid nitrogen (-196°C) and maintained in that state while X-ray diffraction data were collected. For B_2H_6 , the crystallography needed even colder temperatures, requiring liquid helium (-269°C), an approach that

had never before been attempted.

Lipscomb and his colleagues discovered that these molecules had unprecedented cage-like shapes quite unlike hydrocarbons, and they developed a detailed theory to explain why. The central idea was the 'two-electron, three-centre bond', in which three atoms are bound together by a single pair of electrons. Although three-centre B–H–B bonds had been postulated earlier by others, Lipscomb's extension of the idea to the B–B–B bond was a major conceptual advance.

This intuitive leap was the key to understanding the borane structures. More broadly, it implied that such multicentre bonding might allow the stable existence of many other types of molecular clusters. Indeed it does, in the form of thousands of known carboranes (clusters containing carbon and boron) and other compounds that incorporate most of the elements in the periodic table, including the 'nonclassical hydrocarbons' such as the pyramidal-shaped $\text{C}_6(\text{CH}_3)_6^{2+}$ ion.

Bill encouraged unconventional thinking even at the risk of occasional error, as he put it in a classic paper in 1954. Those of us who worked with the Colonel (he was delighted to be named a Kentucky Colonel by the state's governor in 1973) came to appreciate his philosophy of the 'intuitive leap' as a method of advancing science.

Throughout his scientific career, music remained a serious avocation. He served as principal clarinetist with the Pasadena and Minneapolis Civic Orchestras, helped to found the New Friends of Chamber Music in Minneapolis and played regularly for years with members of the Boston Symphony Orchestra.

A member of the Baker Street Irregulars (devotees of Sherlock Holmes), Bill was also given to quoting from Arthur Conan Doyle and Lewis Carroll in his papers to make a point. These facets of his personality, as well as the revolution he fomented in the understanding of the covalent bond, form his lasting legacy. ■

Russell N. Grimes is a professor emeritus of chemistry at the University of Virginia. He was a graduate student in the Lipscomb group at the University of Minnesota in 1958–59 and at Harvard University in 1960–62.
e-mail: rng@virginia.edu

E. SEGRE VISUAL ARCHIVES/AMERICAN INST. PHYSICS/SPL

Species loss revisited

Conservationists predict massive extinctions as a result of habitat loss. Habitat loss undoubtedly does drive extinctions, but dealing with an unmet assumption that underlies these predictions yields much lower estimates. [SEE LETTER P.368](#)

CARSTEN RAHBEK & ROBERT K. COLWELL

Scientists generally agree that Earth is facing a biodiversity crisis, losing species 100 to 1,000 times faster than the normal background rate of extinction¹ and resulting in the sixth period of mass extinction in Earth's history. On page 368 of this issue, He and Hubbell² provide a fresh perspective on predictions of the rate of this species loss.

Previous periods of mass extinction were driven by global changes in climate and in atmospheric chemistry, bolide impacts and volcanism³. This time, species extinction is a result of interaction and competition for resources with another species — humans. We are immensely successful. Our numbers are many times higher than ecological theory would predict for a species with our life history and body mass. We explore, populate and drastically alter almost all corners of the Earth and modify the global climate. Loss of habitat is predicted by various studies to cause the extinction of 20–50% of all species in just half a century³. These estimates began to surface decades ago, but sceptics have repeatedly demanded evidence of widespread extinction, asking 'Where are the bodies?'. If proof is not forthcoming, they argue, then politicians and decision-makers should denounce the biodiversity crisis as a myth⁴.

He and Hubbell² question the way that extinction rates attributed to habitat loss have most often been estimated. Biologists have struggled for decades to estimate how many species are going extinct. Traditionally, the answer has relied on estimates based on an almost universal ecological relationship — when we inventory the species in an area of natural habitat, the list grows as the area is increased. Using theoretical or empirically derived functions to describe this species–area relationship (SAR), it has been assumed that, by working backwards along the SAR, one can estimate the number of species that would be lost to extinction if a larger area were reduced by habitat loss.

A classic rule of thumb says that if habitat area is reduced by 90% (comparable to actual habitat loss in many regions), roughly one-half of its species will be lost. He and Hubbell cite studies using SAR that predicted the loss of 50% of all species by the year 2000 — predictions

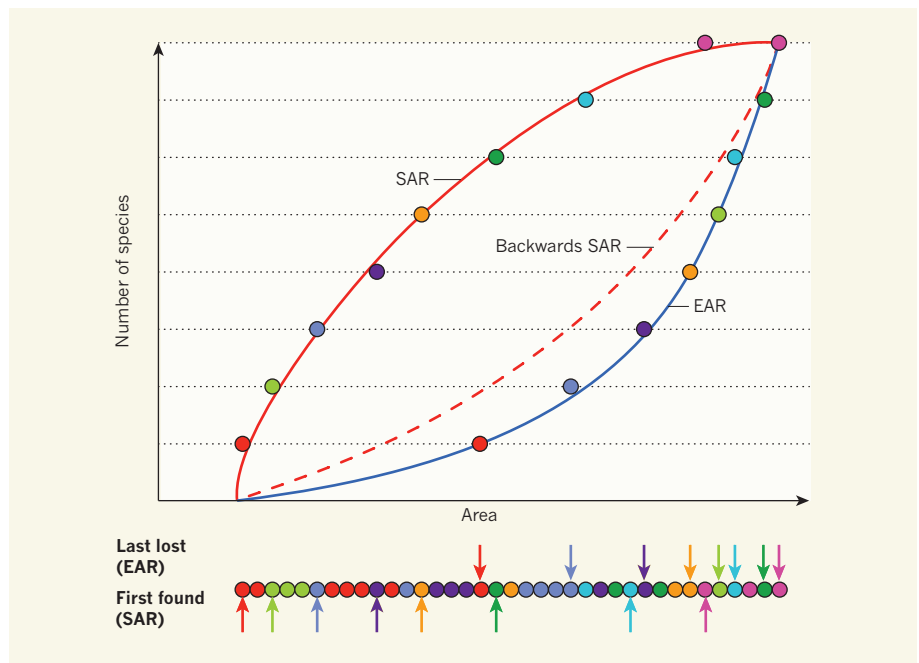


Figure 1 | Estimating species extinctions due to habitat loss. This hypothetical example shows the contrast between use of the backwards species–area relationship (SAR), traditionally used to predict extinctions, and the true endemics–area relationship (EAR) that correctly estimates extinctions with increasing area lost. The coloured circles under the graph represent the spatial ordering of 37 individuals (each occupying one unit of area) of 8 species along a transect through a habitat, each species indicated by a different colour. The total area surveyed increases with each individual encountered. As the first individual of each species is found, the SAR rises by one species, whereas the EAR is incremented only when the last individual of a species is accounted for along the transect. The backwards SAR mirrors the loss of species as area is reduced by moving right-to-left along the SAR. He and Hubbell² demonstrate mathematically and with examples for trees and birds that, for realistic (aggregated) spatial patterns of individuals and species, the backwards SAR always lies above the true EAR, thus overestimating expected rates of extinction. Species aggregation is simulated here by placing dots of the same colour closer to one another than expected at random.

that clearly have not been fulfilled. The discrepancy is well known and has often been explained as 'extinction debt', a time-lag before populations reduced in numbers by habitat loss actually become extinct. Individuals of long-lived species may continue to reproduce or simply live on without reproducing, even if the current living space for the species cannot sustain viable populations over time.

The authors² explain why this traditional 'backwards' use of SAR is fundamentally flawed for typical spatial diversity patterns, and show that this approach can produce drastic overestimation of extinction rates.

The problem with the traditional approach

is surprisingly simple. With increasing habitat area, the SAR rises by one species unit each time the first individual of a species new to the inventory is encountered (Fig. 1). Additional individuals of a species already encountered add nothing to the species count. By contrast, with decreasing habitat area an extinction does not occur until the last individual of a species is encountered. The authors show that, for the aggregated spatial patterns characteristic of species in real communities, the predicted number of extinctions rises more gradually with increasing habitat loss than predicted by the 'backwards' SAR (Fig. 1). The curve that correctly describes the rate of extinction as habitat

area decreases is called the endemics–area relationship (EAR). This was proposed more than a decade ago by Harte and Kinzig⁵ and, they persuasively argued⁶, is more appropriate than the SAR for estimating species extinctions, especially under non-random spatial distributions⁷. (A species is endemic if it is found only within some specified area.)

In their novel conceptualization of the problem, He and Hubbell² show that both the classic SAR and the EAR can be derived from a sampling theory based on spatially explicit patterns of individuals. Applying this approach to empirical data for woody plants in the rain-forest and North American birds, which show typical patterns of spatial aggregation, they quantify the substantial discrepancy between backwards-SAR-based and EAR-based extinction rate predictions (finding overestimation as high as 160% for the plants). Importantly, the authors also justify the use of a simple approximation for the EAR that is robust to variation in species' spatial patterns and scale.

He and Hubbell, then, strongly question the use of SAR to estimate extinction rates not only from direct habitat loss, but also from projected species-range contractions expected to occur under climate change (see ref. 8 for an example). But they emphasize that their results do not in any way imply that there is not an ongoing mass extinction of species, nor that extinction debt is not a genuine biological phenomenon. Even with a better way to estimate rates of future species extinctions, there is still a need to obtain the data required to use the EAR to make more rigorous estimates. There is also the daunting problem of rigorously inferring extinction — showing that the last individual of a species has indeed died.

We invest heavily in infrastructure to store and make accessible the data we have, but by and large we have all but halted investment in discovering and describing the diversity of species with which we share the Earth. At best we have described only about 10% of all living multicellular species. If we 'fog' a tropical tree, literally hundreds of insect species unknown to science fall to the ground. Every year, many new species of even the best-known groups, the mammals and birds, are described. For only a fraction of the known species do we have even a rough idea of their entire geographical distributions.

Most of Earth's biodiversity occurs in tropical regions where species occur at low density and tend to have tiny geographical ranges. The first individual of such a species encountered in a brief inventory is not far from the last to go when extinction threatens, compared with populous, widespread species at higher latitudes. Thus, when modifying tropical habitat through forestry, mining or agriculture, we rarely have an idea which species inhabit the environment we are about to affect, nor the exact consequences of our action. The 'body bags' are rarely counted. ■

Carsten Rahbek is at the Center for Macroecology, Climate and Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. **Robert K. Colwell** is in the Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut 06269-3043, USA. e-mails: crabek@bio.ku.dk; robert.colwell@uconn.edu

ASTRONOMY

Bound and unbound planets abound

Two teams searching for extrasolar planets have jointly discovered a new population of objects: ten Jupiter-mass planets far from their host stars, or perhaps even floating freely through the Milky Way. [SEE LETTER P.349](#)

JOACHIM WAMBSGANSS

Two decades ago, we had no idea whether planets orbiting stars other than the Sun existed at all. Today, more than 500 exoplanets have been discovered, and the field of exoplanet research has advanced to become one of the most captivating branches of astronomy. Observational techniques now aim to address questions such as what the atmosphere and weather are like on some of these planets, and to determine their global statistical properties. On page 349 of this issue, the MOA and OGLE research teams¹ provide an exciting result for exoplanetary science: the discovery of a population of planets that have roughly the mass of Jupiter and separations from their putative host stars of at least ten times Earth's distance to the Sun.

The teams' finding¹ is based on gravitational microlensing, an established technique for detecting exoplanets that is well placed for statistical studies of exoplanets. There are two particularly exciting aspects to the discovery of this new exoplanetary population. The first is the authors' conclusion that, on average, there is more than one Jupiter-mass planet per Milky Way star. The second is the evidence that these planetary-mass objects could be at great distances from their host stars. Some of them could even be floating freely through the Milky Way — that is, they might not be gravitationally bound to any star at all.

Gravitational microlensing is one of a suite of planet-search techniques. The methods are truly complementary to one another, each probing different planetary properties and having its own particular strengths². But most of them detect and explore nearby exoplanets. By contrast, microlensing probes more distant

1. Pimm, S. L., Russell, G. J., Gittleman, J. L. & Brooks, T. M. *Science* **269**, 347–350 (1995).
2. He, F. & Hubbell, S. P. *Nature* **473**, 368–371 (2011).
3. Barnosky, A. D. et al. *Nature* **471**, 51–57 (2011).
4. Lomborg, B. *The Skeptical Environmentalist: Measuring the Real State of the World* (Cambridge Univ. Press, 2001).
5. Harte, J. & Kinzig, A. P. *Oikos* **80**, 417–427 (1997).
6. Kinzig, A. P. & Harte, J. *Ecology* **81**, 3305–3311 (2000).
7. Green, J. L. & Ostling, A. *Ecology* **84**, 3090–3097 (2003).
8. Thomas, C. D. et al. *Nature* **427**, 145–148 (2004).

planets, using the host star–planet system as a magnifying glass. When a foreground star (the lens) passes in front of a distant, background star, the latter is magnified and displays a characteristic 'light curve'³. The two observables that characterize such a microlensing event are the height of the light curve's magnification peak and the duration of the magnification, which depends, among other parameters, on the mass of the lens: the lower the mass, the shorter the duration. Originally proposed as a way of searching for dark matter, it soon became clear that microlensing could also be used to detect planetary systems⁴: a planet orbiting the foreground star would produce a secondary peak in the light curve (Fig. 1).

Microlensing offers two advantages over other methods: it has the potential to yield the most representative statistical sample of Milky Way planets and it is, in principle, sensitive enough to detect Earth-mass objects^{5,6} with current technology. However, the downside is that microlensing events are rare: fewer than one in a million stars in the central part of the Milky Way are microlensed at any given time by a foreground lensing star. And even if every such lensing star had a Jupiter-mass planet at a few times the Earth–Sun distance, only about 1% of these planets would be detected, owing to the exact geometric alignment required between the background star, the planetary system and an observer on Earth. So discovering such microlensing events is akin to finding a needle in a haystack.

To tackle these statistical challenges, a handful of independent research teams have developed advanced techniques to monitor the brightness of about 100 million Milky Way stars every few days. These techniques have allowed the teams to routinely find about

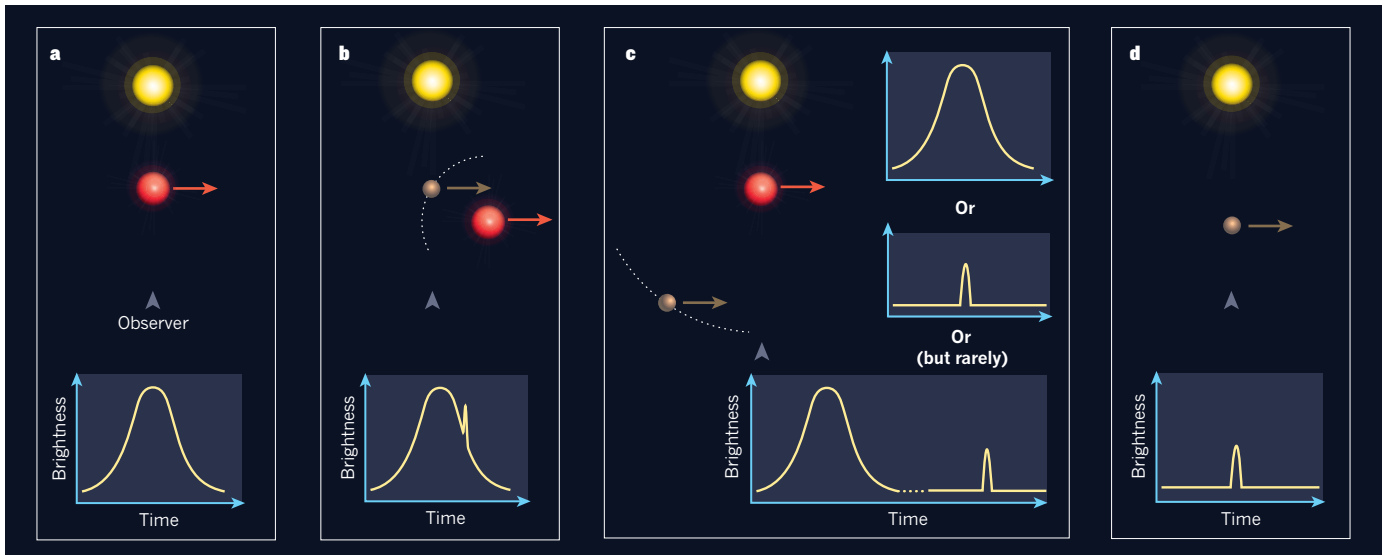


Figure 1 | Planet microlensing. **a**, When a foreground star (red) passes in front of a distant, background star (yellow), it bends the background star's light and causes it to brighten and fade with a characteristic 'light curve'. **b**, For a foreground system composed of a star and an orbiting planet (brown) that are close to each other, the brightening and fading can be accompanied by a sharp secondary peak due to the planet. **c**, If the host star and planet are far

apart, most observed light curves will display either the broad peak associated with the star or the sharp peak associated with the planet; very rarely will the alignment between the background star, the foreground planetary system and the observer on Earth be such that both peaks are observed (the two observations can be years apart). **d**, For an isolated planet without a host star, the observed light curve will always display a single, short-duration peak.

1,000 (stellar) microlensing events per year. So far, however, only about a dozen exoplanets have been detected by microlensing. Nevertheless, impressive results have been derived on the abundance of planets in the Milky Way: planetary systems similar to our own are expected around one sixth of all stars⁷, and cold Neptune-mass planets are common⁸.

In a specially designed study, the MOA (Microlensing Observations in Astrophysics) team⁹ monitored 50 million Milky Way stars for about two years, each at least once per hour. In this way, they were able to detect microlensing events of very short duration. In a careful analysis of the data — which excludes all known sources of contamination — the team has now discovered¹ 474 individual microlensing events, ten of which lasted for less than two days. The researchers then added independent data obtained by the OGLE (Optical Gravitational Lensing Experiment) team¹⁰, to substantiate their original conclusions that there are many more short-duration microlensing events than expected from the known population of stars and brown dwarfs in the Milky Way. The authors¹ interpret this over-abundance of short events as being produced by a thus-far unknown population of Jupiter-mass objects.

Because the observed light curves for the ten very short-duration microlensing events do not show any signature of a possible host star, the authors¹ conclude that these Jupiter-mass objects must be located at distances from their host stars of at least ten times the Earth–Sun distance. When comparing their derived abundance of Jupiter-mass objects with upper limits on abundances of wide-separation exoplanets from direct detections, they¹ argue that it is very likely that most of their newly discovered

planetary-mass objects are unbound. These conclusions prompt at least two questions.

To be or not to be called a planet — that is the first (linguistic) question. After the first discovery, about a decade ago, of isolated low-mass objects in young star-forming regions¹¹, a heated discussion ignited over what to call these entities. Among the contending denominations were 'free-floating planets', 'isolated planetary-mass objects', 'objects formerly called planets' and 'rogue planets'. One of the contentious issues is whether the mass and the dynamic state of the objects concerned alone should determine their class name, or whether their formation history should also be considered. The International Astronomical Union (IAU) succeeded, in 2006, in re-defining what a planet is. But it postponed the definition of an exoplanet. In light of the discovery of a probable new class of objects¹, it may now be worthwhile to reconsider these definitions^{12,13}.

To be or not to be a bound planet — that is the second (astronomical) question. If these objects do turn out to be unbound, we want to understand how they reached this state. The MOA and OGLE teams provide¹ plausible arguments, but various hypotheses for the formation and dynamic state of the objects seem possible, and certainly deserve further investigation. Ultimately, the question of whether these objects are bound to stars or freely floating through the Milky Way will be answered through astronomical observations. In the former case, the relative motion between the background star and the foreground star–planet system will occasionally be oriented such that the background star will be magnified a second time by the focusing effect of the planet's host star¹⁴. This

second (broader) peak may well happen a few years after (or before) the planetary blip in the light curve. Another signature of a bound planet, known as astrometric microlensing, is a minute change in the position of the background star during the magnification¹⁵.

The implications of this discovery¹ are profound. We have a first glimpse of a new population of planetary-mass objects in our Galaxy. Now we need to explore their properties, distribution, dynamic states and history. A continuation of high-cadence ground-based microlensing observations will surely shed some further light on these objects. But dedicated observations by satellite telescopes with large viewing angles will be pivotal for a full understanding of this population. Well-developed concepts for such projects^{16–18} on both sides of the Atlantic guarantee a head start. Exploring unbound (former) satellites of stars with bound (future) satellite telescopes of planet Earth will open up a new chapter in the history of the Milky Way. ■

Joachim Wambsganss is at the Zentrum für Astronomie der Universität Heidelberg, Astronomisches Rechen-Institut, 69120 Heidelberg, Germany.
e-mail: jkw@uni-hd.de

1. The Microlensing Observations in Astrophysics (MOA) Collaboration & The Optical Gravitational Lensing Experiment (OGLE) Collaboration. *Nature* **473**, 349–352 (2011).
2. Perryman, M. *The Exoplanet Handbook* (Cambridge Univ. Press, 2011).
3. Paczyński, B. *Astrophys. J.* **304**, 1–5 (1986).
4. Mao, S. & Paczyński, B. *Astrophys. J.* **374**, L37–L40 (1991).
5. Beaulieu, J.-P. et al. *Nature* **439**, 437–440 (2006).
6. Liebig, C. & Wambsganss, J. *Astron. Astrophys.* **520**, A68 (2010).

7. Gould, A. *et al. Astrophys. J.* **720**, 1073–1089 (2010).
8. Sumi, T. *et al. Astrophys. J.* **710**, 1641–1653 (2010).
9. www.phys.canterbury.ac.nz/moa/
10. <http://ogle.astrouw.edu.pl>
11. Zapatero Osorio, M. R. *et al. Science* **290**, 103–107 (2000).
12. www.iau.org/public_press/news/detail/iau0603
13. Soter, S. *Astron. J.* **132**, 2513–2519 (2006).
14. Di Stefano, R. & Mao, S. *Astrophys. J.* **457**, 93–101 (1996).
15. Han, C. *Astrophys. J.* **644**, 1232–1236 (2006).
16. <http://sci.esa.int/euclid>
17. <http://wfirst.gsfc.nasa.gov>
18. Bennett, D. P. *et al. http://arxiv.org/abs/1012.4486*

ECONOMICS

A positive side of disaster

In October 1998, a hurricane visited death and destruction on Honduras, with flooding and mudslides. A case history of a rural community documents how recovery from that event produced socio-economic improvement.

ARUN AGRAWAL

It may seem heartless, or at least inappropriate, to talk of natural disasters as windows of opportunity. Because they are major shocks to socio-economic systems, disasters such as Cyclone Nargis (Myanmar, 2008) or Hurricane Katrina (United States, 2005) are capable of wiping out entire settlements, destroying the lives and livelihoods of thousands, if not hundreds of thousands, and wreaking material destruction on a massive scale¹. The poor suffer most². But writing in *Proceedings of the National Academy of Sciences*, McSweeney and Coomes³ suggest that even such disastrous events may sometimes yield positive economic outcomes for the rural poor.

Four years after the devastation caused by Hurricane Mitch in Honduras in 1998 (Fig. 1), the authors found that the indigenous Tawahka community of Krausirpi, in the northeast of the country, was better off than before the

disaster. Furthermore, income and assets in the community were more equitably distributed. Households had on average three times more land; poorer groups and women had gained more land; agricultural production had been re-established; sources of income were more diverse; a new land-tenure system was in place; and the community was probably more resilient to future climate shocks.

What explains this relatively positive picture? Two processes helped: education among those who had been land-poor, and new norms of land tenure rooted in the diffuse character of decision-making among the Tawahka. Younger families with higher education found it easier to access new employment and wage opportunities offered by non-governmental organizations and state agencies; these same families tended to be smaller, and had previously found it difficult to clear land. Krausirpi is a land-surplus community; after Mitch, local residents quietly laid claim to new areas

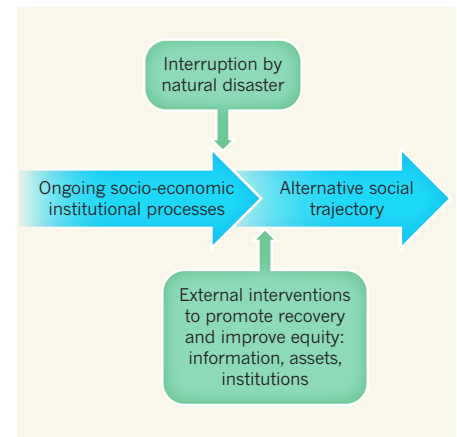


Figure 2 | Disasters, interventions and outcomes. The appropriate application and withdrawal of interventions can alter the social trajectory of a hard-hit community.

of arable land through individual negotiations — no explicit, centralized decision was made to adopt a different system of land allocation.

McSweeney and Coomes, then, conclude that Hurricane Mitch effectively reset the social, economic and institutional machinery in Krausirpi. The hurricane forced all households to look for new land and, in the process, to accept a new mechanism through which to allocate land. It encouraged local residents to look for new sources of income and access to assets. Finally, it permitted households to take advantage of existing human capital in novel ways because of wage opportunities that became available with the presence of development organizations.

What does this example have to say to those interested in post-disaster reconstruction — to those many decision-makers and aid workers who want not only to provide immediate succour to victims of disasters but also to do so in a way that makes recovery sustainable and equitable⁴? The first point to highlight is that disasters are a natural 'reset button' — what happens in their wake is shaped by historical forces, to be sure, but they also enable greater leveraging power to new resources, fresh endeavours and innovative institutions, because older structures and processes lose at least part of their historical force.

A second point is that it is possible to improve both incomes and equity in the wake of disasters. But doing so requires a focus not only on productive opportunities, but also on strategies that favour the poor and the less powerful; a focus on the creation of new income streams and assets, but through access strategies that are equitably distributed; and a focus on processes of institutional change, but those in which less advantaged households and groups have a voice. In each case, part of the attention is on enhancing incomes, but also, and importantly, on achieving an equitable outcome. Figure 2 summarizes these points.

Finally, many analysts of disaster relief and



Figure 1 | Honduras, 2 November 1998. Hurricane Mitch wreaked havoc.

reconstruction have observed the tension between the imperative to provide immediate relief and efforts to launch self-organizing processes of change that sustain recovery and equity⁵. Immediate disaster relief often requires quick decisions, has only limited opportunities for participatory interventions, and is typically externally driven. Long-term positive changes, on the other hand, require local buy-in and activation of a community's capacities both through strategic interventions and through disengagement at the appropriate time. Successing internal institutional change

that is structured by the interests of weaker and poorer groups is a crucial precursor to disengagement. Knowing when to restrict or cease to provide external material support, and instead facilitate a transition to educational and institutional support mechanisms that recognize local capacities and provide opportunities for the less powerful, is necessary for disaster relief to be effective in the long run. ■

Arun Agrawal is in the School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michigan 48103, USA.

e-mail: arunagra@umich.edu

1. World Bank/United Nations. *Natural Hazards, UnNatural Disasters: The Economics of Effective Prevention* (World Bank, 2010).
2. World Bank. *Building Resilient Communities: Risk Management and Response to Natural Disasters through Social Funds and Community-Driven Development Operations* (World Bank, 2009).
3. McSweeney, K. & Coomes, O. *Proc. Natl Acad. Sci. USA* **108**, 5203–5208 (2011).
4. Pelling, M. *Natural Disasters and Development in a Globalizing World* (Routledge, 2003).
5. Arnold, M. & Burton, C. *Protecting and Empowering Vulnerable Groups in Recovery* (World Bank, in the press).

EARTH SCIENCE

A deep foundry

Melting and solidification of iron alloys in Earth's core may explain structural complexity in the solid inner core, and alter the way we think about the dynamics of the deep interior. SEE LETTER P.361

BRUCE BUFFETT

Textbooks depict Earth as having an onion-layered structure with a solid steel ball at the centre. The central body, known as the inner core, is thought to have formed by gradual cooling and solidification of the surrounding liquid outer core¹. On page 361 of this issue, Gubbins and colleagues² turn convention on its head by arguing that a large fraction of the inner core's surface is melting. Our understanding of both the structure and the dynamics of the core may change as a consequence.

The authors' conclusion is based on a numerical model³ that simulates convection and magnetic-field generation in the liquid core. Cooling of the liquid core drives convection, but it is the more massive and sluggish mantle surrounding it that regulates the rate of cooling. Spatial variations in heat flux at the top of the core exert a strong influence on the pattern of fluid flow⁴. In the authors' simulations, cold fluid is focused into narrow plumes, which descend to the inner-core boundary and promote localized solidification. Elsewhere, a broad return flow is associated with warm fluid that persistently exceeds the melting temperature at the inner-core boundary.

Temperatures in the fluid's interior can exceed the boundary temperature because the core is mainly cooled from above rather than heated from below. Fluid parcels become warmer relative to a decreasing background temperature if the parcels are not cooled at the average rate. Cooling produces net

growth of the inner core, and is so intense that solidification is required below narrow regions of cold fluid in order to offset large areas of melting in warmer regions.

Fractionation of impurities in the liquid on solidification is expected to enrich the solid in iron⁵; therefore, melting should produce a dense liquid that pools on top of the inner core. Gubbins *et al.*² argue that such a melt layer offers a simple explanation for unusual values recorded for seismic velocities near the

surface of the inner core⁶. In addition, heterogeneity in grain size and composition at that surface may eventually be buried by inner-core growth, possibly explaining the complex structure detected inside the inner core⁷. All in all, melting of the inner core provides a tidy explanation for several observations, although a few details remain to be explored.

The authors pay careful attention to several problems that arise in applying their model to Earth. Two additional points are worth mentioning. The first is the perennial concern about the validity of numerical models, given that the physical parameters are very far from realistic values. More specifically, could small-scale turbulence, largely absent from the current models, disperse cold plumes long before they reach the inner-core boundary? As the numerical models improve we can expect to gain better insight into their reliability.

A second question involves the role of composition in the melting temperature.

Impurities in the liquid core depress the liquidus temperature relative to that of pure iron by 600 kelvin or more⁸. An iron-rich melt is expected to solidify before a liquid with the bulk composition of the outer core. So what happens after the inner core melts?

We expect the inner-core boundary to represent the top of a mushy region where solid and liquid coexist⁹. A temperature increase at the boundary initially promotes melting. However, a small amount of melt enriches the surrounding liquid in iron, which elevates the local liquidus temperature and brings the interface back into equilibrium (Fig. 1). Given the magnitude of the melting-point depression, a small amount of melt should be sufficient to compensate for thermal fluctuations in the liquid outer core. Would small variations in melt volume be detectable in seismic observations? This question will require a better understanding of the phase diagram.

The work of Gubbins and colleagues² opens a door onto new enquiries. Melting and solidification

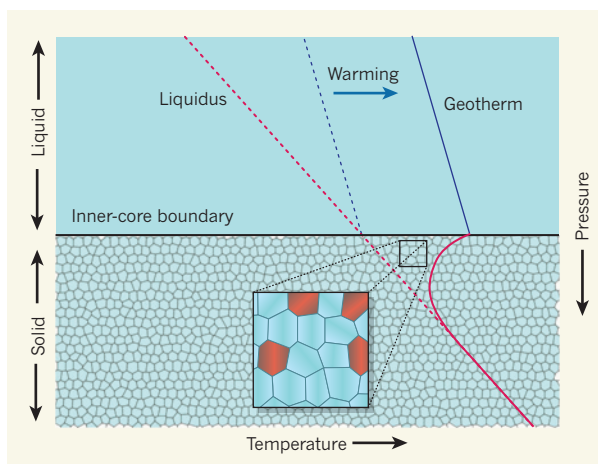


Figure 1 | Melting at the inner-core boundary. Gubbins *et al.* suggest² that warming in the liquid outer core produces widespread melting at the boundary with the solid inner core. In this phase diagram, the initial position of the inner-core boundary is defined by the intersection of the core temperature (geotherm; dotted blue line) and the melting temperature (liquidus; dotted red line). An increase in the geotherm (solid blue line) promotes melting of the iron-rich solid, diluting the concentration of impurities in the liquid. That raises the liquidus temperature (solid red line) until it intersects the warmer geotherm, re-establishing thermodynamic equilibrium. The dense melt resides within the porous solid near the top of the inner core, as shown in the inset.

of the inner core allows greater interaction with the surrounding liquid core, and raises the possibility that surprising phenomena are yet to be discovered. Recent speculations^{10,11} about a steady translational motion of the inner core demonstrate that strange things are possible. The final chapter of this story is yet to be written. ■

Bruce Buffett is in the Department of Earth & Planetary Science, University of California, Berkeley, Berkeley, California 94720-4767, USA. e-mail: bbuffett@berkeley.edu

1. Jacobs, J. A. *Nature* **172**, 297–298 (1953).

2. Gubbins, D., Sreenivasan, B., Mound, J. & Rost, S. *Nature* **473**, 361–363 (2011).
3. Sreenivasan, B. & Jones, C. A. *Geophys. J. Int.* **164**, 467–476 (2006).
4. Olson, P. & Christensen, U. R. *Geophys. J. Int.* **151**, 809–823 (2002).
5. Alfè, D., Gillan, M. J. & Price, G. D. *Contemp. Phys.* **48**, 63–80 (2007).
6. Souriau, A. & Poupinet, G. *Geophys. Res. Lett.* **18**, 2023–2026 (1991).
7. Sun, X. & Song, X. *Earth Planet. Sci. Lett.* **269**, 56–65 (2008).
8. Gillan, M. J., Alfè, D., Brodholt, J., Vočadlo, L. & Price, G. D. *Rep. Prog. Phys.* **69**, 2365–2441 (2006).
9. Fearn, D. R., Loper, D. E. & Roberts, P. H. *Nature* **292**, 232–233 (1981).
10. Monnereau, M. *et al. Science* **328**, 1014–1017 (2010).
11. Alboussière, T., Deguen, R. & Melzani, M. *Nature* **466**, 744–747 (2010).

EPIGENETICS

Tet proteins in the limelight

Tet proteins mediate the hydroxymethylation of DNA. New work reveals their function in gene regulation and the extent of their activity throughout the genome of embryonic stem cells. SEE ARTICLE P.343 & LETTERS P.389, P.394 & P.398

**NATHALIE VÉRON &
ANTOINE H. F. M. PETERS**

During mammalian development, the one-cell zygote gives rise to a multitude of cell types. This remarkable process is controlled by protein machines that interpret the genetic code and regulate the expression of genes, in part by chemically modifying chromatin (DNA–protein complexes). One such modification is the addition of a methyl group at the 5-position of the cytosine base in DNA (5mC) — an alteration that serves a crucial role in the epigenetic (or cell-to-cell) inheritance of gene expression during development. However, proteins of the Tet enzyme family can modify this DNA mark further by hydroxylating the methyl group to form 5-hydroxymethylcytosine (5hmC)^{1–3}. Five papers^{4–8}, including four in this issue, report on the extent of 5hmC modification across the genome of mouse embryonic stem cells and on the role of Tet proteins in regulating gene expression.

The 5mC modification is required for genome stability and thus the embryo's viability. It is also needed for the repression of genes and repetitive genomic sequences; for X-chromosome inactivation; and for genomic imprinting (in which, for some genes, either the maternal or the paternal copy is expressed). Classical studies revealed that 5mC is erased in primordial germ cells and during early embryo development, and that this process occurs independently of DNA replication as

the cells divide⁹. What's more, genes containing 5mC can become active in differentiated cells, supporting the notion of active demethylation^{10,11}. This idea, along with researchers' ability to epigenetically reprogram cells (either by the technique of somatic-cell nuclear transfer or by induced pluripotency experiments), inspired the search for factors that mediate 5mC demethylation, although initially there was limited success¹².

Because proteins of the Tet family (Tet1–3) can convert 5mC to 5hmC, they have been considered promising candidates for mediating DNA demethylation. But this novel enzymatic means of demethylation leads to obvious questions. Where in the genome do Tet proteins bind? How do they affect the stability/turnover of 5mC? Does this influence gene expression during the cell cycle and in development? Do Tet proteins alleviate gene silencing by converting 5mC to 5hmC, or do they protect against aberrant *de novo* methylation, thereby preventing silencing? And finally, how is 5hmC processed further? The latest studies^{4–8} shed light on these issues.

Williams *et al.* (page 343)⁴, Wu *et al.*⁸ and Ficiz *et al.* (page 398)⁷ localized 5hmC in the genome of mouse embryonic stem (ES) cells using methods that predominantly recognize DNA sequences bearing multiple 5hmC marks. Pastor and colleagues (page 394)⁶ developed two alternative methods that possess increased sensitivity for single 5hmCs. The general finding is that 5hmC levels across the genome are low. Nonetheless, the mark is

significantly enriched at CpG dinucleotides within genes, particularly at exons (protein-coding regions)^{4,6,8}, and this is correlated positively with gene-expression levels⁸, as reported before¹³.

The authors^{4,7} further investigated 5hmC localization in ES cells deficient in the three enzymes that are involved in *de novo* DNA methylation and in the maintenance of this modification (TKO cells). Like 5mC, 5hmC is absent in these cells. Hence, 5hmC probably arises from the processing of pre-existing 5mC along the gene body during transcription.

In mammalian genomes, CpG dinucleotides occur at roughly 60% of promoter sequences, forming 'CpG islands' (CGIs) that are largely devoid of 5mC. Remarkably, Williams *et al.*⁴ and Wu *et al.*⁸ report that 5hmC is significantly enriched only at a fraction of such CGIs, predominantly at inactive promoters, many of which are marked with two modifications to chromatin: histone 3 lysine 4 trimethylation (H3K4me3), which activates genes, and histone 3 lysine 27 trimethylation (H3K27me3), which is repressive. By contrast, Williams and colleagues⁴ and Wu *et al.* (page 389)⁵ show that Tet1 occupies not just double-marked (bivalent) repressed promoters, but also active promoters marked with H3K4me3 alone. There seems to be a strong correlation between Tet1 occupancy and CpG density at CGIs and at other genomic regions⁴.

Most 5hmC-positive promoters display low levels of 5mC⁴. Does this mean that Tet1 is involved in the turnover of pre-existing 5mC, for instance at bivalent promoters, to protect against inappropriate *de novo* DNA methylation? If so, 5hmC might be a transient intermediate of active demethylation, potentially triggering downstream processing pathways such as the base-excision repair mechanism^{10,14}. The varying 5hmC levels among CGIs (such as at H3K4me3/Tet1 promoters compared with H3K4me3/H3K27me3 promoters) raises questions about the occurrence and kinetics of 5mC and 5hmC turnover as a function of transcriptional activity and/or local chromatin configuration.

Curiously, reducing Tet1 expression in normal ES cells resulted in only a minor increase in 5mC levels at CGIs and along the gene body^{4,5}. This may be because a substantial fraction of 5hmC sites is devoid of Tet1⁴. It might further reflect technical limitations in reducing the active Tet1 enzyme and/or partial functional redundancy between Tet1 and Tet2, which is also expressed in ES cells^{7,15}. The true relevance of Tet enzymes to 5mC turnover therefore remains to be determined.

Even more surprisingly, the experiments^{4,5} reveal that Tet1 predominantly has repressive, rather than activating, functions on its direct target genes. Many of the genomic regions that are inactivated by Tet1 are occupied by proteins of the Polycomb repressive complex 2 (PRC2), which catalyses the formation of

H3K27me3 and represses the transcription of genes involved in specifying cellular identity during development. Although the authors^{4,5} did not detect any direct biochemical interaction between Tet1 and PRC2 proteins, Tet1 depletion directly or indirectly alleviated transcriptional repression and PRC2 recruitment at genes to which both proteins were bound⁵.

In addition, Tet1 binds directly to Sin3A, a co-repressor protein essential for inhibiting the transcription of a subset of genes that are also repressed by Tet1⁴. By analysing the expression of selected Tet1 target genes, Williams *et al.* further show in TKO cells (which lack the Tet1 substrate 5mC) that, on Tet1 depletion, these genes are misregulated in a similar way to control cells⁴. So, at least in ES cells, Tet1 seems to be required for transcriptional gene regulation independently of its enzymatic activity, and possibly for regulating the recruitment of proteins that define chromatin states at CGIs.

Several strands of evidence point to a functional role for 5mC/5hmC turnover in differentiation. During ES-cell differentiation, Tet1 and Tet2 levels decrease, as do 5hmC levels, whereas 5mC levels increase, concomitantly with changes in gene expression^{7,15}. Moreover, around one-third of genomic regions containing 5hmC in ES cells acquire 5mC during development⁴. And 5hmC levels are almost ten times higher in cells of the brain's cerebellum region that have stopped dividing than in

proliferating ES cells¹³. Finally, Tet1 regulates gene expression induced by neuronal activity and 5mC turnover in the dentate gyrus of the adult mouse brain¹⁰. To clarify the role of Tet proteins in active demethylation, genetic approaches aiming at delineating their catalytic and non-catalytic functions are required, particularly in non-dividing cells and during development. And so the search for factors indispensable to DNA demethylation continues. Stay tuned. ■

Nathalie Véron and Antoine H. F. M. Peters
are at the Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland.
e-mail: antoine.peters@fmi.ch

1. Kriaucionis, S. & Heintz, N. *Science* **324**, 929–930 (2009).
2. Tahiliani, M. *et al.* *Science* **324**, 930–935 (2009).
3. Ito, S. *et al.* *Nature* **466**, 1129–1133 (2010).
4. Williams, K. *et al.* *Nature* **473**, 343–348 (2011).
5. Wu, H. *et al.* *Nature* **473**, 389–393 (2011).
6. Pastor, W. A. *et al.* *Nature* **473**, 394–397 (2011).
7. Ficiz, G. *et al.* *Nature* **473**, 398–402 (2011).
8. Wu, H. *et al.* *Genes Dev.* **25**, 679–684 (2011).
9. Reik, W. *Nature* **447**, 425–432 (2007).
10. Guo, J. U., Su, Y., Zhong, C., Ming, G. & Song, H. *Cell* **145**, 423–434 (2011).
11. Klug, M. *et al.* *Genome Biol.* **11**, R63 (2010).
12. Ooi, S. K. T. & Bestor, T. H. *Cell* **133**, 1145–1148 (2008).
13. Song, C.-X. *et al.* *Nature Biotechnol.* **29**, 68–72 (2011).
14. Hajkova, P. *et al.* *Science* **329**, 78–82 (2010).
15. Koh, K. P. *et al.* *Cell Stem Cell* **8**, 200–213 (2011).

PLANT BIOLOGY

Unveiling the Casparian strip

The Casparian strip in plant roots is a diffusion barrier that directs water and solutes from the soil to the water-conducting tissues. Proteins involved in making the strip have at long last been identified. [SEE LETTER P.380](#)

MARKUS GREBE

The roots of land plants are usually in direct contact with the soil, which provides water, nutrients and other solutes but may also contain toxic substances. To control the uptake of molecules from the soil, vascular plants are thought to have evolved the protective cell layer that surrounds their water-conducting system. This layer, called the endodermis, acts as an inner skin¹. A distinguishing mark of endodermal cells is the local thickening of their transverse cell walls, which was first described in 1865 by Robert Caspary² and later named the Casparian strip. On page 380 of this issue, Roppolo and colleagues³ now describe a family of proteins that are precisely located in the plasma membrane

adjacent to the Casparian strip and are needed for its formation.

The Casparian strip contains the polymeric molecules lignin and suberin, which locally impregnate the transverse cell walls of endodermal cells (Fig. 1a)⁴. This impregnation hinders diffusion of water and solutes through the cell-wall space^{1,5}, thus forcing them to pass through the outer membrane (plasma membrane) and the interior (cytoplasm) of the endodermal cells that act as a selective barrier (Fig. 1b). Although regulators of asymmetric cell division and cell-fate determination during endodermis formation have been identified^{6,7}, proteins specifically needed to make the Casparian strip have remained unknown.

Roppolo *et al.*³ reasoned that messenger RNAs encoding proteins specifically

involved in Casparian strip formation should be expressed in the endodermis and should include transcripts for membrane proteins. They therefore analysed mRNA transcripts strongly expressed in the root endodermis of the model plant *Arabidopsis thaliana* and identified five mRNAs encoding similar proteins that are predicted to contain four membrane-spanning regions. These proteins belonged to a large, uncharacterized plant-specific protein family.

When the five proteins were fused to green fluorescent protein (GFP) and expressed in *Arabidopsis* plants, they were all found in the endodermis. Intriguingly, the proteins precisely localized to the plasma-membrane domain at the site of Casparian strip formation, the Casparian strip domain (CSD), and were therefore named Casparian strip proteins 1 to 5 (CASP1–5). Roppolo *et al.* found that CASP1 and CASP3 act together during Casparian strip formation, because the strip remained patchy and not smoothly assembled in plants in which the *casp1* and *casp3* genes had both been disrupted. CASP1 and CASP3 may even interact in a protein complex, because the proteins immunoprecipitated together and in complexes with other CASPs. Whether this involves direct interaction in CASP heteromers, and what role the other CASPs play, remain topics for future studies.

Clearly, CASP1 and CASP3 are needed for the correct formation of the Casparian strip. But at what stage do they act? And are they structural components, or do they instruct strip assembly? Clues to the answers came from images of the distribution of the CASP1–GFP fusion protein in living roots. Roppolo *et al.*³ observed that localization of CASP1 at the CSD preceded establishment of the diffusion barrier between the outer and inner cell walls of endodermal cells, thus suggesting an early role for CASPs during Casparian strip formation. Strikingly, CASP1 was uniformly distributed within the plasma membrane during early cell differentiation, but it then relocated and became concentrated at the future CSD. Here, CASP1 first occurred at the plasma membrane in small patches that then progressively fused.

Up to this stage, the distribution of CASP1 within the cell was sensitive to the inhibition of a pathway — the endocytic trafficking pathway — that is needed for the uptake of proteins from the plasma membrane and for their redistribution within the cell. This implied that CASP1 is, at least in part, targeted from an overall distribution to its specific localization at the CSD by endocytic recycling. Intriguingly, once this specific localization was established, CASP1 became immobile, as if strongly attached to the cell wall.

To see whether CASPs could instruct the formation of a specific plasma-membrane domain, Roppolo *et al.*³ expressed CASPs in other tissues. This resulted in an even accumulation of CASPs at the plasma membrane, or,

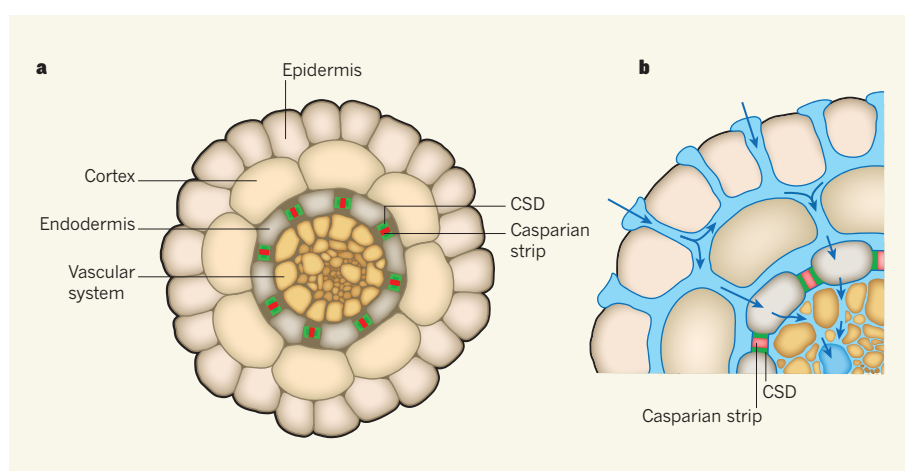


Figure 1 | The Casparian strip. **a**, Schematic cross-section through an *Arabidopsis* root. The outermost cell layer (epidermis) surrounds the cortical cell layer, which in turn surrounds the endodermis. The transverse cell walls of endodermal cells that connect cell walls of cortical and vascular cells are impregnated by a cell-wall thickening, the Casparian strip (red). The plasma-membrane domain of differentiated endodermal cells in direct contact with the strip is called the Casparian strip domain (CSD) (green). It is the localization site of the previously unknown Casparian strip proteins (CASPs) identified by Roppolo and colleagues³. **b**, Close-up of the upper-left quarter of the root cross-section, showing the flow of water and solutes (blue) through the cell-wall space. Because of impregnation by the Casparian strip, water and solutes cannot cross transverse cell walls and are redirected to pass through the plasma membrane and the interior of endodermal cells. These act as selective diffusion barriers, before releasing water and selected solutes into the vascular system.

in the case of CASP5, at internal cellular membranes. These findings indicate that additional endodermis-intrinsic factors are needed to direct CASPs to a specific plasma-membrane domain and that CASP5 differs from other family members.

It remains to be seen whether CASPs themselves instruct Casparian strip formation, but they are certainly the first proteins known to specifically participate in this process. Moreover, their early redistribution to the CSD by endocytic recycling provides a mechanistic insight into the formation of this domain. These observations open the door for studies on how specific CASP targeting to the CSD is achieved, and what mechanisms underlie subsequent CASP immobilization.

The availability of *casp* mutants should now allow the question of what the Casparian strip actually does to be addressed. Is it really important for the plant's protection against abiotic and biotic stress factors¹? How do *casp* mutants cope with boron toxicity, for example? Boron is an essential nutrient, but it is toxic for plants at high concentrations. Proteins regulating boron influx and efflux across the plasma membrane can be found in the outer and inner endodermal plasma membrane, respectively^{8,9}. These transporters are excluded from the CSD^{3,8}, where the CASPs are located. Thus, it will be interesting to see whether CASPs help to create a barrier that regulates boron flux.

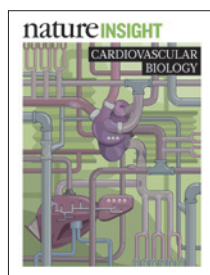
Roppolo *et al.*³ propose that the endodermis is functionally analogous to an animal epithelium, and point out that, with CASPs, plants have evidently invented their own way of making an equivalent to a diffusion barrier

between animal epithelial cells — the tight junction. Indeed, tight junctions contain proteins, the claudins¹⁰, that do not obviously resemble CASPs at the sequence level but that are similar in their size and overall organization into four transmembrane domains. Although claudins are essential for the formation of diffusion barriers in animals¹⁰, Roppolo *et al.*³ did not observe defective diffusion of a tracer through the malformed Casparian strip of plants lacking CASP1 and CASP3, probably because other CASPs could compensate for potential defects.

Roppolo *et al.* have started to reveal how plants build the Casparian strip, and it will be exciting to learn how it acquires its properties as a diffusion barrier. We can look forward to further unveiling of the mechanisms of CASP action and interaction in the near future. ■

Markus Grebe is in the Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, SE-90187 Umeå, Sweden. e-mail: markus.grebe@plantphys.umu.se

1. van Fleet, D. S. *Bot. Rev.* **27**, 165–220 (1961).
2. Caspary, R. *Jb. Wissensch. Bot.* **4**, 101–124 (1865/66).
3. Roppolo, D. *et al. Nature* **473**, 380–383 (2011).
4. Zeier, J., Ruel, K., Ryser, U. & Schreiber, L. *Planta* **209**, 1–12 (1999).
5. Robards, A. W. & Robb, M. E. *Science* **178**, 980–982 (1972).
6. Di Laurenzio, L. *et al. Cell* **86**, 423–433 (1996).
7. Helariutta, Y. *et al. Cell* **101**, 555–567 (2000).
8. Alassimone, J., Naseer, S. & Geldner, N. *Proc. Natl Acad. Sci. USA* **107**, 5214–5219 (2010).
9. Takano, J. *et al. Proc. Natl Acad. Sci. USA* **107**, 5220–5225 (2010).
10. Furuse, M. & Tsukita, S. *Trends Cell Biol.* **16**, 181–188 (2006).



Cover illustration by
Nik Spencer

Editor, *Nature*
Philip Campbell

Publishing
Nick Campbell

Insights Editor
Ursula Weiss

Production Editor
Nicola Bailey

Senior Art Editor
Kelly Buckheit Krause

Art Editor
Nik Spencer

Sponsorship
Gerard Preston

Production
Emilia Orviss

Marketing
Elena Woodstock,
Hannah Phipps

Editorial Assistant
Hazel Mayhew

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel: +44 (0) 20 7833 4000
e: nature@nature.com



nature publishing group

A healthy vasculature is crucial to our survival: blood vessels act as conduits to deliver oxygen and nutrients to every tissue in our body, remove waste and allow immune surveillance. Vascular endothelial cells also contribute to processes such as haematopoiesis and organ development during embryogenesis. Not surprisingly, vascular dysfunction is linked to diverse disorders, from cancer to eye diseases, and can also trigger heart failure and death. Understanding how vessels grow and function has huge potential for improving human health, and this exciting topic is developing at a rapid pace.

Anti-angiogenic agents that block vascular endothelial growth factor (VEGF) are being used in the clinic to treat patients with cancer and eye diseases, but efficacy has been limited. Peter Carmeliet and Rakesh Jain review the frontline research into the mechanisms regulating normal and pathophysiological angiogenic growth: they discuss the clinical experience with VEGF blockers and describe recently identified signalling pathways that contribute to vessel growth, maturation and quiescence, which may provide new avenues to improve anti-angiogenic therapy.

Mark Lindsay and Harry Dietz explore the pathogenesis of aortic aneurysm; an enlargement and weakening of the arteries, which can lead to fatal tearing. They show how research has implicated the cytokine transforming growth factor- β in the pathogenesis of this disorder, and how this has led to new therapeutic opportunities with losartan, an angiotensin II receptor antagonist.

Atherosclerosis is a disease of the arterial wall, in which lipid-filled plaques develop in the inner lining of the artery. Preclinical research has led to many hypotheses about the pathogenesis of plaque formation. Peter Libby, Paul Ridker and Göran Hansson discuss the limitations of animal models of atherosclerosis and the difficulties with extrapolating these findings to human disease, and suggest better ways to consolidate preclinical and clinical research.

A potentially fatal consequence of ruptured atherosclerotic plaques is myocardial infarction, which causes a massive loss of cardiomyocytes, leading to heart dysfunction or complete heart failure. In the last review of this series, Michael Laflamme and Charles Murry discuss the emerging prospects for regenerating the damaged myocardium, a dynamic area of research that draws from advances in stem-cell biology, developmental biology and tissue engineering.

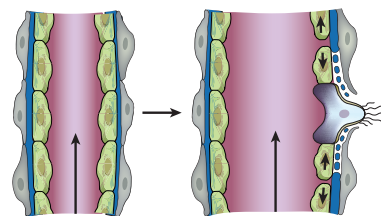
Clare Thomas
Senior Editor

CONTENTS

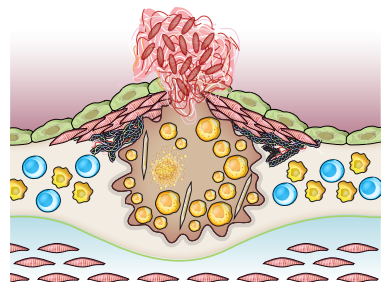
REVIEWS

298 Molecular mechanisms and clinical applications of angiogenesis
Peter Carmeliet & Rakesh K Jain

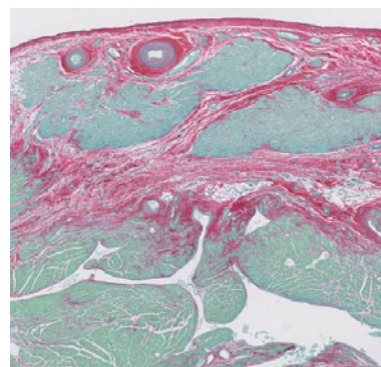
308 Lessons on the pathogenesis of aneurysm from heritable conditions
Mark E Lindsay & Harry C Dietz



317 Progress and challenges in translating the biology of atherosclerosis
Peter Libby, Paul M Ridker & Göran K Hansson



326 Heart regeneration
Michael A Laflamme & Charles E Murry



Molecular mechanisms and clinical applications of angiogenesis

Peter Carmeliet^{1,2} & Rakesh K. Jain³

Blood vessels deliver oxygen and nutrients to every part of the body, but also nourish diseases such as cancer. Over the past decade, our understanding of the molecular mechanisms of angiogenesis (blood vessel growth) has increased at an explosive rate and has led to the approval of anti-angiogenic drugs for cancer and eye diseases. So far, hundreds of thousands of patients have benefited from blockers of the angiogenic protein vascular endothelial growth factor, but limited efficacy and resistance remain outstanding problems. Recent preclinical and clinical studies have shown new molecular targets and principles, which may provide avenues for improving the therapeutic benefit from anti-angiogenic strategies.

Blood vessels arose in evolution to allow haematopoietic cells to patrol the organism for immune surveillance, to supply oxygen and nutrients and to dispose of waste. Vessels also produce instructive signals for organogenesis in a perfusion-independent manner (Box 1). Although beneficial for tissue growth and regeneration, vessels can fuel inflammatory and malignant diseases, and are exploited by tumour cells to metastasize and kill patients with cancer. Because vessels nourish nearly every organ of the body, deviations from normal vessel growth contribute to numerous diseases. To name just a few, insufficient vessel growth or maintenance can lead to stroke, myocardial infarction, ulcerative disorders and neurodegeneration, and abnormal vessel growth or remodelling fuels cancer, inflammatory disorders, pulmonary hypertension and blinding eye diseases^{1,2}.

Several modes of vessel formation have been identified (Fig. 1). In the developing mammalian embryo, angioblasts differentiate into endothelial cells, which assemble into a vascular labyrinth — a process known as vasculogenesis (Fig. 1b). Distinct signals specify arterial or venous differentiation³. Subsequent sprouting ensures expansion of the vascular network, known as angiogenesis (Fig. 1a). Arteriogenesis then occurs, in which endothelial cell channels become covered by pericytes or vascular smooth muscle cells (VSMCs), which provide stability and control perfusion. Tissues can also become vascularized by other mechanisms, but the relevance of these processes is not well understood. For example, pre-existing vessels can split by a process known as intussusception, giving rise to daughter vessels (Fig. 1c). In other cases, vessel co-option occurs, in which tumour cells hijack the existing vasculature (Fig. 1d), or tumour cells can line vessels — a phenomenon known as vascular mimicry (Fig. 1e). Putative cancer stem-like cells can even generate tumour endothelium⁴ (Fig. 1f). Although debated, the repair of healthy adult vessels or the expansion of pathological vessels can be aided by the recruitment of bone-marrow-derived cells (BMDCs) and/or endothelial progenitor cells to the vascular wall. The progenitor cells then become incorporated into the endothelial lining in a process known as postnatal vasculogenesis. Collateral vessels, which bring bulk flow to ischaemic tissues during revascularization, enlarge in size by distinct mechanisms, such as the attraction and activation of myeloid cells⁵.

The revascularization of ischaemic tissues would benefit millions, but therapeutic angiogenesis remains an unmet medical need. Instead, more success has been achieved by targeting the vascular supply in cancer and eye diseases⁶. In this Review, we describe key molecular targets in

angiogenesis and discuss the clinical experience with the most widely used class of anti-angiogenic agent — blockers of vascular endothelial growth factor (VEGF, also known as vascular permeability factor or VPF). Rather than providing an encyclopaedic survey, we focus on some of the recently discovered mechanisms and principles, and on targets with translational potential.

Vessel branching, maturation and quiescence

We first provide the current view of the sequential steps of vessel branching (quiescence, activation and resolution), before discussing the molecular players involved in more depth (Fig. 2). In a healthy adult, quiescent endothelial cells have long half-lives and are protected against insults by the autocrine action of maintenance signals such as VEGF, NOTCH, angiopoietin-1 (ANG-1) and fibroblast growth factors (FGFs). Because vessels supply oxygen, endothelial cells are equipped with oxygen sensors and hypoxia-inducible factors — such as prolyl hydroxylase domain 2 (PHD2) and hypoxia-inducible factor-2 α (HIF-2 α), respectively — which allow the vessels to re-adjust their shape to optimize blood flow. Quiescent endothelial cells form a monolayer of phalanx cells with a streamlined surface, interconnected by junctional molecules such as VE-cadherin and claudins. These endothelial cells are ensheathed by pericytes, which suppress endothelial cell proliferation and release cell-survival signals such as VEGF and ANG-1. Endothelial cells and pericytes at rest produce a common basement membrane.

When a quiescent vessel senses an angiogenic signal, such as VEGF, VEGF-C, ANG-2, FGFs or chemokines, released by a hypoxic, inflammatory or tumour cell, pericytes first detach from the vessel wall (in response to ANG-2) and liberate themselves from the basement membrane by proteolytic degradation, which is mediated by matrix metalloproteinases (MMPs) (Fig. 2a). Endothelial cells loosen their junctions, and the nascent vessel dilates. VEGF increases the permeability of the endothelial cell layer, causing plasma proteins to extravasate and lay down a provisional extracellular matrix (ECM) scaffold. In response to integrin signalling, endothelial cells migrate onto this ECM surface. Proteases liberate angiogenic molecules stored in the ECM such as VEGF and FGF, and remodel the ECM into an angio-competent milieu. To build a perfused tube and prevent endothelial cells from moving en masse towards the angiogenic signal, one endothelial cell, known as the tip cell, becomes selected to lead the tip in the presence of factors such as VEGF receptors, neuropilins (NRPs) and the NOTCH ligands

¹Laboratory of Angiogenesis and Neurovascular Link, Vesalius Research Center, VIB, Leuven B-3000, Belgium. ²Laboratory of Angiogenesis and Neurovascular Link, Vesalius Research Center, K. U. Leuven, Leuven B-3000, Belgium. ³Edwin L. Steele Laboratory for Tumor Biology, Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA.

BOX 1

Perfusion-independent role of endothelial cells

During embryogenesis, the invasion of endothelial cells into nascent organs confers inductive signals to promote organogenesis, even in the absence of blood flow. This suggests that endothelial cells not only form passive conduits for delivering oxygen but also establish organ-specific vascular niches, which stimulate organogenesis by the production of paracrine-tropic 'angiocrine' factors⁹². Endothelial cells show remarkable heterogeneity in different organs. These organ-specific endothelial cells release signals for pancreatic differentiation, reconstitution of haematopoietic stem cells and

expansion of neuronal precursors, and give rise to haematopoietic progenitors by endothelial-to-haematopoietic transition. The vascular adventitia — the outer layer of vessels — also hosts vessel-resident stem and progenitor cells. Emerging evidence indicates that such perfusion-independent activities of endothelial cells also promote tumorigenesis⁹². In addition to constituting the building blocks of vessels and delivering nutrients and oxygen, tumour endothelial cells allow the recruitment of pro-angiogenic bone-marrow-derived cells.

DLL4 and JAGGED1 (Fig. 2a). The neighbours of the tip cell assume subsidiary positions as stalk cells, which divide to elongate the stalk (stimulated by NOTCH, NOTCH-regulated ankyrin repeat protein (NRARP), WNTs, placental growth factor (PlGF) and FGFs) and establish the lumen (mediated by VE-cadherin, CD34, sialomucins, VEGF and hedgehog) (Fig. 2b). Tip cells are equipped with filopodia to sense environmental guidance cues such as ephrins and semaphorins, whereas stalk cells release molecules such as EGFL7 into the ECM to convey spatial information about the position of their neighbours, so that the stalk elongates. A hypoxia-inducible program, driven by HIF-1 α , renders endothelial cells responsive to angiogenic signals. Myeloid bridge cells aid fusion with another vessel branch, allowing the initiation of blood flow. For a vessel to become functional, it must become mature and stable. Endothelial cells resume their quiescent phalanx state (Fig. 2c), and signals such as platelet-derived growth factor B (PDGF-B), ANG-1, transforming growth factor- β (TGF- β), ephrin-B2 and NOTCH cause the cells to become covered by pericytes. Protease inhibitors known as tissue inhibitors of metalloproteinases (TIMPs) and plasminogen activator inhibitor-1 (PAI-1) cause the deposition of a basement membrane, and junctions are re-established to ensure optimal flow distribution. Vessels regress if they are unable to become perfused.

The VEGF family

Given the complexity of a process such as angiogenesis, it is remarkable that a single growth factor, VEGF, regulates this process so predominantly. The VEGF family consists of only a few members and distinguishes itself from other angiogenic superfamilies by the largely non-redundant roles of its members. VEGF (also known as VEGF-A) is the main component, and it stimulates angiogenesis in health and disease by signalling through VEGF receptor-2 (VEGFR-2, also known as FLK1)^{7,8}. Neuropilins such as NRP1 and NRP2 are VEGF co-receptors, which enhance the activity of VEGFR-2, but also signal independently⁹. Similar to VEGFR-2 deficiency, the loss of VEGF aborts vascular development². In response to a VEGF gradient, established by soluble and matrix-bound isoforms, tip cells upregulate DLL4 expression, which activates NOTCH in stalk cells; this downregulates VEGFR-2 expression, rendering stalk cells less responsive to VEGF, thereby ensuring that the tip cell takes the lead¹⁰. Soluble VEGF isoforms promote vessel enlargement, whereas matrix-bound isoforms stimulate branching. Paracrine VEGF, released by tumour, myeloid or other stromal cells, increases vessel branching and renders tumour vessels abnormal¹¹, whereas autocrine VEGF, released by endothelial cells, maintains vascular homeostasis¹². Emerging evidence indicates that the biological effect of VEGFR-2 signalling depends on its subcellular localization — for example, for VEGF to induce arterial morphogenesis, VEGFR-2 must signal from intracellular compartments¹³. Activating *VEGFR2* mutations cause vascular tumours, and genetic polymorphisms in *VEGF* and/or its receptors co-determine pathological angiogenesis^{14,15}, whereas the blockade of VEGF signalling can target angiogenic vessels in malignant and ocular disease in humans. VEGF protein or gene transfer stimulates

vessel growth in ischaemic tissues, but often in association with undesired leakage and vessel abnormalities.

VEGF-C, a ligand of the VEGFR-2 and VEGFR-3 receptors, activates blood-vessel tip cells¹⁶. VEGFR-3 is necessary for the formation of the blood vasculature during early embryogenesis, but later becomes a key regulator of lymphangiogenesis — the formation of new lymphatic vessels from pre-existing ones¹⁷. In zebrafish, in which the first embryonic vein arises by segregation of venous-fated endothelial cells from a common precursor vessel, the sprouting of venous endothelial cells is restricted by VEGFR-2 but promoted by VEGFR-3 (ref. 18). Venous-derived angiogenesis in the arterial trunk also relies on VEGFR-3 signalling. Anti-VEGFR-3 antibodies that inhibit receptor dimerization or ligand binding slow down tumour growth synergistically, and enhance the inhibition of tumour growth by VEGFR-2 blockade, making VEGFR-3 another anti-angiogenic candidate¹⁶.

Originally discovered as a VEGF homologue, PlGF was also expected to be an angiogenic factor. However, unlike VEGF, PlGF is dispensable for development and is relevant only in disease^{19,20}. PlGF is a multitasking cytokine that stimulates angiogenesis by direct or indirect mechanisms, and also activates bone-marrow-derived endothelial progenitor and myeloid cells, as well as stromal cells, to create a nurturing 'soil' for tumour cells, in addition to activating tumour cells¹⁹. By skewing the polarization of tumour-associated macrophages (TAMs), the loss of PlGF improves vessel perfusion and maturation, and enhances responses to chemotherapy²¹. PlGF blockade by neutralizing anti-PlGF antibodies phenocopies the anti-angiogenic effects of genetic *Plgf* (also known as *Pgf*) deficiency in spontaneous mouse tumour models and diseases such as ocular neovascularization²². Yet other PlGF-blocking strategies fail to inhibit the growth of tumours in transplantable tumour models²³. The therapeutic potential of PlGF blockade in patients with cancer thus remains to be established. In preclinical models, PlGF protein or gene delivery increases the revascularization of ischaemic tissues.

Deficiency of the VEGF family member VEGF-B in mice does not impair angiogenesis in normal development, and cannot compensate for VEGF blockade after birth¹⁹. VEGF-B has only restricted angiogenic activity in certain tissues such as the heart, yet it promotes neuronal survival and induces metabolic effects^{19,24}. Divergent effects of VEGF-B on pathological angiogenesis have been reported, and it has been shown to promote the growth of cardiac vessels, without inducing adverse effects such as increased permeability or leakage²⁵.

The precise role of the VEGFR-1 receptor (also known as FLT-1) in angiogenesis remains elusive^{19,26}. VEGFR-1 exists both as a membrane-anchored signalling-competent form and as a soluble secreted form (also known as sFLT-1). By trapping its ligands, sFLT-1 can assist the guidance of the emerging branch or inhibit sprouting altogether. Because of its weak tyrosine kinase activity, VEGFR-1 may act as a decoy for VEGF, moderating the amount of free VEGF available to activate VEGFR-2 and explaining why VEGFR-1 loss results in vessel overgrowth¹⁹. However, intracellular VEGFR-1 signalling in angiogenic endothelial, stromal and myeloid cells stimulates pathological

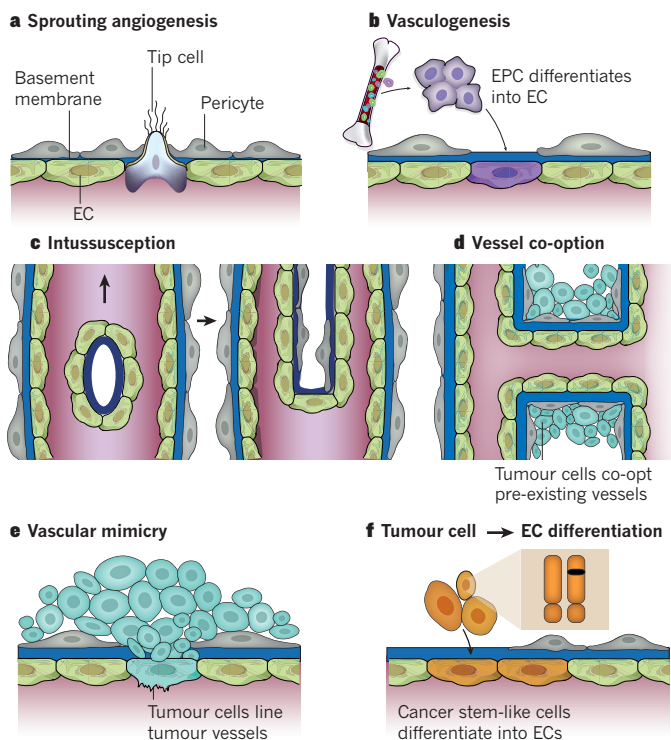


Figure 1 | Modes of vessel formation. There are several known methods of blood vessel formation in normal tissues and tumours. **a–c**, Vessel formation can occur by sprouting angiogenesis (**a**), by the recruitment of bone-marrow-derived and/or vascular-wall-resident endothelial progenitor cells (EPCs) that differentiate into endothelial cells (ECs; **b**), or by a process of vessel splitting known as intussusception (**c**). **d–f**, Tumour cells can co-opt pre-existing vessels (**d**), or tumour vessels can be lined by tumour cells (vascular mimicry; **e**) or by endothelial cells, with cytogenetic abnormalities in their chromosomes, derived from putative cancer stem cells (**f**). Unlike normal tissues, which use sprouting angiogenesis, vasculogenesis and intussusception (**a–c**), tumours can use all six modes of vessel formation (**a–f**).

angiogenesis²⁶. VEGFR-1 signalling also promotes the growth of VEGFR-1⁺ tumour cells in response to autocrine VEGF production in an angiogenesis-independent manner²⁷, and upregulates MMP9 in endothelial cells at the premetastatic site. There is evidence to suggest that VEGFR-1⁺ haematopoietic progenitors form a premetastatic niche in distant organs, but this finding is debated^{28,29}. Neutralizing anti-PlGF, anti-VEGFR-1 and anti-VEGFR-2 antibodies are in early clinical development.

The PDGF family

For vessels to function properly, they must be mature and covered by mural cells. Several growth-factor families, such as PDGFs, angiopoietins and TGF- β , contribute to this process³⁰. To stabilize endothelial cell channels, angiogenic endothelial cells release PDGF-B to chemoattract PDGF receptor- β (PDGFR- β)⁺ pericytes^{31,32}. Hence, pericyte deficiency after PDGF-B ablation causes vessel leakage, tortuosity, microaneurysm formation and bleeding. Knockout of the genes encoding the PDGF-B protein retention motif (necessary for pericyte adhesion) in mice results in tumour vessel fragility and hyperdilation, whereas PDGFR- β -hypomorph mice have insufficient pericytes around brain vessels, leading to blood–brain barrier (BBB) defects and neurodegenerative damage owing to the leakage of toxic substances³³. Tumour-derived PDGF-B also recruits pericytes indirectly by upregulating stromal-cell-derived factor-1 α (SDF-1 α ; encoded by *CXCL12*). Besides a local origin, pericytes can also arise from perivascular PDGFR- β ⁺ pericyte progenitors, recruited from the bone marrow³⁴. By inhibiting PDGFR- β signalling in mural cells, VEGF reduces pericyte coverage and renders tumour vessels abnormal.

PDGFR inhibition diminishes tumour growth by causing pericyte

detachment, leading to immature vessels that are prone to regression³⁵. Other pericyte-deficient mouse strains that lack the proteoglycan NG2 (also known as CSPG4) also form abnormal tumour vessels and smaller tumours. Paradoxically, the overexpression of PDGF-B in mice inhibits tumour growth by promoting pericyte recruitment and inducing endothelial cell growth arrest³⁶. Because the survival of endothelial cells depends on pericyte VEGF production, pericytes protect endothelial cells from VEGF withdrawal and confer resistance to VEGF blockade. This protection requires a close endothelial-cell–pericyte interaction, as PDGF-B blockade reduces pericyte coverage and vessel number only when VEGF is produced by pericytes and not by more distant tumour cells³⁷. Initial studies using multi-target receptor tyrosine kinase inhibitors (TKIs) showed that blocking PDGF-B renders mature vessels more sensitive to VEGF blockade by depleting the vessels of pericytes³¹. Recent studies with more specific inhibitors have shown that combination therapy is no more efficient than anti-VEGF monotherapy³⁸.

PDGFR- β ⁺ pericytes have a dual role in metastasis. In primary tumours, pericytes limit tumour cell intravasation, because the more loosely assembled vessel wall is no longer a barrier for disseminating tumour cells after depletion of pericytes³⁹. The absence of pericytes around vessels also correlates with metastasis in patients, and a trial evaluating PDGF-B blockade was aborted because of excessive leakage. These studies indicate that blocking vessel maturation can promote malignancy. However, other reports have shown that pericytes, co-opted by tumour cells at micrometastatic sites, allow tumour colonization by releasing angiogenic factors. Overall, future studies are needed to explore the benefits and risks of PDGF blockade for the treatment of cancer.

PDGF-B blockade may be used therapeutically for non-malignant vascular diseases such as pulmonary hypertension, whereas PDGF-B activation may offer therapeutic opportunities for stabilizing vascular malformations⁴⁰. PDGF-CC, another family member released by cancer-associated fibroblasts in VEGF-inhibitor-resistant tumours, stimulates vessel growth and maturation, and attenuates the response to anti-VEGF treatment^{41,42}. By preventing the activation of perivascular PDGFR- α ⁺ astrocytes, which together with pericytes constitute the BBB, the blockade of PDGF-CC preserves the integrity of the BBB during stroke. Inhibition of PDGF-DD suppresses ocular neovascularization, whereas PDGF-DD overexpression normalizes tumour vessels and improves drug delivery.

TGF- β signalling

Human hereditary haemorrhagic telangiectasia is characterized by vascular malformations. Human genetic studies have shown that this disorder is due to mutations in the genes that encode endoglin (*ENG*) or activin receptor-like kinase (*ALK1*, also known as *ACVRL1*) — receptors of the TGF- β family. Mouse studies have confirmed that the loss of the TGF- β receptors ALK-1, TGFR-1 (also known as ALK-5), TGFR-2 or *ENG* results in arteriovenous malformations, reminiscent of those seen in patients with hereditary haemorrhagic telangiectasia⁴³. However, understanding the molecular basis of this pathway has been challenging owing to inconsistent results. This is partly due to the context-dependent pro- and anti-angiogenic effects of TGF- β family members. Furthermore, although TGF- β promotes VSMC differentiation, and deficiency of *ENG* or *ALK-1* impairs mural cell development, it remains unclear whether other TGF- β components mediate their vascular effects *in vivo* by means of endothelial cells or VSMCs⁴³. Preclinical studies have shown that antibodies against *ENG* or *ALK-1* can inhibit tumour angiogenesis and growth. Several TGF- β blockers are now in early-phase clinical trials.

The FGF superfamily

The superfamily of FGFs and their receptors controls a wide range of biological functions⁴⁴. bFGF was among the first discovered angiogenic factors and, like FGF1, has angiogenic and arteriogenic properties; FGF9 stimulates angiogenesis in bone repair. FGFs activate receptors (FGFRs) on endothelial cells or indirectly stimulate angiogenesis by inducing the release of angiogenic factors from other cell types⁴⁴. For instance, in the heart, FGF-mediated signalling fuels vessel growth by stimulating

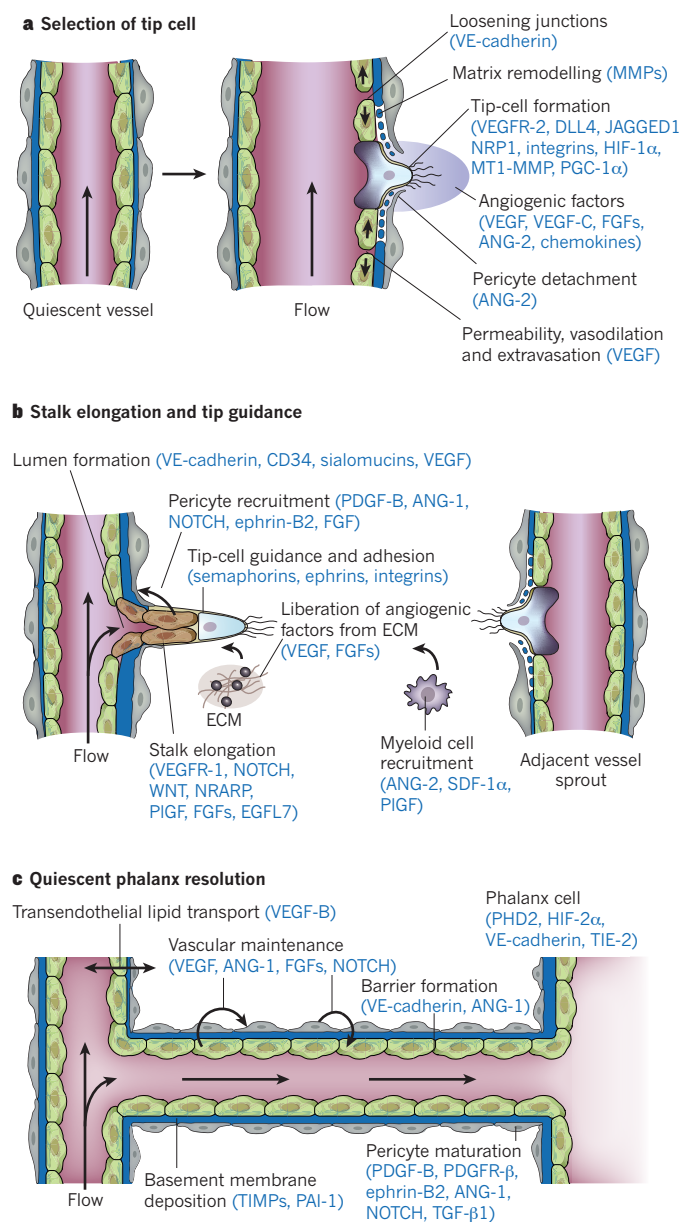


Figure 2 | Molecular basis of vessel branching. The consecutive steps of blood vessel branching are shown, with the key molecular players involved denoted in parentheses. **a**, After stimulation with angiogenic factors, the quiescent vessel dilates and an endothelial cell tip cell is selected (DLL4 and JAGGED1) to ensure branch formation. Tip-cell formation requires degradation of the basement membrane, pericyte detachment and loosening of endothelial cell junctions. Increased permeability permits extravasation of plasma proteins (such as fibrinogen and fibronectin) to deposit a provisional matrix layer, and proteases remodel pre-existing interstitial matrix, all enabling cell migration. For simplicity, only the basement membrane between endothelial cells and pericytes is depicted, but in reality, both pericytes and endothelial cells are embedded in this basement membrane. **b**, Tip cells navigate in response to guidance signals (such as semaphorins and ephrins) and adhere to the extracellular matrix (mediated by integrins) to migrate. Stalk cells behind the tip cell proliferate, elongate and form a lumen, and sprouts fuse to establish a perfused neovessel. Proliferating stalk cells attract pericytes and deposit basement membranes to become stabilized. Recruited myeloid cells such as tumour-associated macrophages (TAMs) and TIE-2-expressing monocytes (TEMs) can produce pro-angiogenic factors or proteolytically liberate angiogenic growth factors from the ECM. **c**, After fusion of neighbouring branches, lumen formation allows perfusion of the neovessel, which resumes quiescence by promoting a phalanx phenotype, re-establishment of junctions, deposition of basement membrane, maturation of pericytes and production of vascular maintenance signals. Other factors promote transendothelial lipid transport.

the release of hedgehog, ANG-2 and VEGF-B. Low levels of FGF are required for the maintenance of vascular integrity, as inhibition of FGFR signalling in quiescent endothelial cells causes vessel disintegration⁴⁵. Aberrant FGF signalling promotes tumour angiogenesis and mediates the escape of tumour vascularization from VEGF- or epidermal growth factor receptor (EGFR)-inhibitor treatment⁴⁶. Development of specific FGF or FGFR inhibitors for blocking angiogenesis is lagging behind, partly because *Fgf1* or *Fgf2* deficiency in mice did not produce vascular defects and the FGF superfamily shows substantial redundancy⁴⁴. FGF protein or gene transfer has been tested for therapeutic angiogenesis, but without sustained success in the clinic.

The ANG and TIE signalling system

Healthy vessels must be equipped with mechanisms to maintain quiescence, while remaining able to respond to angiogenic stimuli. The ANG and TIE family is a binary system to offer such a switch. The human ANG family consists of two receptors, TIE-1 and TIE-2, and three ligands, ANG-1, ANG-2 and ANG-4. ANG-1 functions as a TIE-2 agonist, and ANG-2 functions as a competitive ANG-1 antagonist in a context-dependent manner (ANG-4 has not been as well studied, but is thought to act like ANG-1). Because no ligand for TIE-1 has been identified, this orphan receptor may act as a negative regulator of TIE-2, but its precise role remains elusive⁴⁷. ANG-1 is expressed by mural and tumour cells, whereas ANG-2 is released from angiogenic tip cells. In confluent endothelium, ANG-1 induces TIE-2 clustering *in trans* at cell–cell junctions to maintain endothelial cell quiescence⁴⁸. ANG-1 also stimulates mural coverage and basement membrane deposition, thereby promoting vessel tightness. In the presence of angiogenic stimulators, sprouting endothelial cells release ANG-2, which antagonizes ANG-1 and TIE-2 signalling to enhance mural cell detachment, vascular permeability and endothelial cell sprouting⁴⁷. In accordance, *Tie2* (also known as *Tek*) deficiency in mice causes vascular defects, and activating germline and somatic *TIE2* (*TEK*) mutations in humans result in venous malformations. Tumour-derived ANG-2 also promotes angiogenesis by recruiting pro-angiogenic TIE-2-expressing monocytes (TEMs)⁴⁹.

The overall effects of the ANG–TIE system on tumours are context dependent⁴⁷. ANG-1 stimulates tumour growth by promoting endothelial cell survival and vessel maturation, but it also inhibits tumour cell extravasation and maintains the integrity of healthy vessels outside tumours. These conflicting biological activities warrant caution when considering ANG-1 as an anticancer target. Instead, ANG-2 may be a more appealing therapeutic target because it stimulates tumour angiogenesis and recruits pro-angiogenic TEMs, and ANG-2 inhibition promotes vessel regression and normalization⁵⁰. Given that ANG-2 and VEGF cooperatively increase angiogenesis, co-blockade of VEGF and ANG-2 is superior in inhibiting tumour angiogenesis, metastasis and leakage⁵¹. Various agents that block either TIE-2 or ANG-2 are being evaluated in early-phase clinical trials.

The NOTCH and WNT signalling pathway

The vessel-branching model postulates that, in general, tip cells migrate and stalk cells proliferate. Recent studies have implicated NOTCH signalling in this model¹⁰. In response to VEGF, activation of VEGFR-2 upregulates DLL4 expression in tip cells. In neighbouring stalk cells, DLL4 then activates NOTCH, which downregulates VEGFR-2 but upregulates VEGFR-1; thus, the stalk cells become less responsive to the sprouting activity of VEGF but more sensitive to molecules such as PIGF. Overall, DLL4 and NOTCH signalling restricts branching but generates perfused vessels¹⁰. By upregulating PDGFR- β in NOTCH⁺ mural cells, DLL4 in endothelial cells also stimulates vessel maturation. JAGGED1, another NOTCH ligand expressed by stalk cells, promotes tip-cell selection by interfering with the reciprocal DLL4 and NOTCH signalling from the stalk cell to the tip cell⁵². NOTCH signalling in stalk cells is dynamic over time, because it upregulates its own inhibitor, NRARP⁵³.

An unanticipated complexity is that endothelial cells continuously compete for the tip-cell position by fine-tuning their expression of

BOX 2

Guidance signals in angiogenesis

Tip cells sense guidance cues, similar to how axonal growth cones explore their surroundings. It is therefore not surprising that molecules used by navigating axons are evolutionarily conserved, and molecules such as VEGF also guide neuronal cells. Navigating endothelial cells express receptors for axon-guidance cues, including ephrin receptors (EPH); neuropilins (NRPs) and PLEXIN-D1, which bind semaphorins; ROBO4, which binds slit proteins; and UNC5B, which binds netrin proteins. Given the size and complexity of these families (and the existing controversies⁹³), we illustrate this concept with a few examples that have therapeutic potential.

EPH receptors and their ligands, the ephrins, regulate cell-contact-dependent patterning and can generate bidirectional signals. The signalling cascade in ephrin-expressing cells is known as reverse signalling, whereas signalling in EPH-receptor-expressing cells is termed forward signalling. Ephrin-B2 and its receptor EPHB4 regulate vessel morphogenesis by several mechanisms^{93,94}. During vasculogenesis, the vascular plexus is marked by ephrin-B2⁺ arterial and EPHB4⁺ venous territories. By avoiding repulsive actions, ephrin-B2⁺ and EPHB4⁺ cells prevent intermingling and segregate from each other. In zebrafish, the emigration of venous-fated cells from a precursor vessel leads to segregation of ephrin-B2⁺ arterial endothelial cells in the dorsal aorta and EPHB4⁺ venous endothelial cells in the cardinal vein¹⁸. Moreover, reverse signalling by ephrin-B2 in tip cells induces VEGFR-2 internalization, which is necessary for downstream signalling of VEGFR-2 to elicit VEGF-induced tip-cell filopodial extension⁹⁵. Ephrin-B2 also promotes the recruitment of mural cells and bone-marrow-derived endothelial progenitor cells. In tumours, the overall effect of EPHB4 is

pro-angiogenic, making it a target for anti-angiogenic therapy. Indeed, upregulation of EPHB4 stimulates tumour angiogenesis, whereas EPHB4 blockade has the opposite effect. Other EPH receptors and ephrin ligands, such as EPHA2 and ephrin-A1, have a role in vessel growth and maturation⁹⁴. Notably, ephrin-A1 levels are upregulated in tumours treated with VEGF blockers, suggesting that it contributes to resistance against VEGF blockade⁴⁶. Various therapeutics that target EPH receptors and ephrin ligands are being developed, but the complexity of this signalling system should be kept in mind.

Semaphorins are secreted or membrane-anchored, and bind to plexin proteins or their NRP co-receptors. The loss of *Plxnd1* in mice induces erroneous navigation of vessels, because endothelial cells cannot recognize the repulsive semaphorin-3E (SEMA3E) signals in their environment. Many semaphorins inhibit tumour angiogenesis, including SEMA3A, SEMA3B, SEMA3D, SEMA3F and SEMA4A, whereas SEMA3C and SEMA4D promote tumour angiogenesis. NRPs bind ligands such as semaphorins and VEGF, but the vascular defects observed in *Nrp1*-deficient embryos are attributable to defective VEGF signalling, rather than defective semaphorin signalling. An antibody that blocks the binding of VEGF, but not of SEMA3A, to NRPs also inhibits tumour angiogenesis. Dual targeting with antibodies that block both VEGF and NRP1 is more effective than single-agent therapy, presumably because the antivasculature remodelling effects of anti-NRP1 antibodies keep vessels in a VEGF-dependent state. In addition, a soluble NRP2B variant with increased VEGF affinity enhances the tumour-growth-inhibitory activity of an antibody that blocks the interactions of VEGF with VEGFR-2 but not with NRPs.

VEGFR-2 versus VEGFR-1, indicating that this signalling circuit is constantly re-evaluated as cells meet new neighbours⁵⁴. In accordance, the inhibition of DLL4 and NOTCH signalling induces the formation of more numerous but hypoperfused vessels, resulting in tumour hypoxia and growth inhibition⁵⁵. However, chronic DLL4 blockade in healthy animals results in vascular neoplasms⁵⁶, and endothelial cell inactivation of RBP-J, a transcription factor downstream of NOTCH, also leads to uncontrolled angiogenesis. Although these data indicate that quiescent phalanx cells need low-level NOTCH signalling, they also warrant caution against the indiscriminate use of DLL4 and NOTCH inhibitors for the treatment of cancer. Signalling by the hedgehog family members also participates in embryonic vasculogenesis, vascular morphogenesis and tube formation, as well as in arterial specification, by regulating NOTCH expression³.

Endothelial cells express various types of WNT ligand and their frizzled (FZD) receptors, of which several stimulate endothelial cell proliferation. NOTCH activates WNT signalling in proliferating stalk cells during vessel branching⁵³, explaining why NOTCH, which usually suppresses proliferation and promotes quiescence, stimulates proliferation of stalk cells *in vivo*. WNT also activates NOTCH in a reciprocal-feedback system, because WNT signals in endothelial cells induce a NOTCH-like phenotype, characterized by branching defects, loss of venous identity and aberrant vascular remodelling⁵⁷. Gene-inactivation of some of the WNT and FZD members in mice (*Wnt2*, *Wnt5a*, *Fzd4* and *Fzd5*) causes vascular defects, whereas the combined loss of *Wnt7a* and *Wnt7b* impairs brain angiogenesis and BBB formation⁵⁸. Because some WNT members inhibit angiogenesis, specific blockers of these proteins will be required.

Integrins and proteases

The ECM provides a physical link between vascular cells and their surrounding tissues. Endothelial cells possess mechanisms to interact with and alter the matrix. Integrins are heterodimeric receptors that mediate adhesion to ECM and immunoglobulin superfamily molecules^{59,60}.

Upregulation of the integrins $\alpha_3\beta_3$ and $\alpha_5\beta_3$ permits growing endothelial cells to bind to provisional matrix proteins in the tumour milieu; these proteins include vitronectin, fibrinogen and fibronectin, both in native and degraded forms. These adhesive interactions provide survival cues and traction for invading endothelial cells. Other integrins involved in angiogenesis include $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_4\beta_1$, $\alpha_5\beta_1$, $\alpha_6\beta_1$, $\alpha_6\beta_1$ and $\alpha_6\beta_4$ (refs 59, 60).

In addition to signalling induced by ligating ECM components, integrins regulate angiogenesis by other mechanisms. Given their ability to interact with several extracellular molecules and transmit signals in a bidirectional manner, integrins function as 'hubs', orchestrating endothelial cell and VSMC behaviour during angiogenesis^{59,61}. Hence, the binding of integrins to growth factors (such as VEGF, FGFs and ANG-1) or their receptors (VEGFR-2 and FGFRs) stimulates vessel growth. Integrins also upregulate and activate zymogen proteases in invading tip cells, and promote vessel maturation by regulating interactions between endothelial cells, pericytes and the basement membrane. Other integrins promote the adhesion of angiogenic BMDCs to tumour endothelial cells. Recent studies have highlighted the complexities in understanding the role of $\alpha_5\beta_3$ in pathological angiogenesis, as tumour angiogenesis was stimulated by gene deficiency in mice but inhibited by pharmacological blockade^{59,60}. Nonetheless, integrin blockers are now being evaluated in the clinic.

Quiescent endothelial cells and pericytes share a common basement membrane, which not only physically restrains these cells but also keeps them quiescent owing to the antiproliferative properties of the ECM components. During branching, proteolytic remodelling of the ECM liberates these cells for unrestricted movement and converts the characteristics of the basement membrane into a pro-angiogenic environment. Distinct proteases such as MMPs modulate angiogenesis by several mechanisms⁶². They promote endothelial cell migration and tube formation by proteolytically remodelling the basement membrane, by executing directional matrix proteolysis (membrane type 1-MMP) or by exposing chemotactic cryptic motifs sites in the ECM. MMPs and

BOX 3

Hypoxia and epigenetic regulation of angiogenesis

The prolyl hydroxylase domain (PHD) proteins PHD1–3 are oxygen-sensing enzymes that hydroxylate the hypoxia-inducible factor (HIF) proteins HIF-1 α and HIF-2 α when sufficient oxygen is available. Once hydroxylated, HIFs are targeted for proteasomal degradation⁹⁶. Under hypoxia, PHDs become inactive, and HIFs initiate broad transcriptional responses to increase the oxygen supply by angiogenesis, through the upregulation of angiogenic factors such as VEGF⁹⁷. HIFs are also activated in non-hypoxic conditions by oncogenes and growth factors, allowing tumour cells to stimulate angiogenesis before they become deprived of oxygen. In general, HIF-1 α promotes vessel sprouting, whereas HIF-2 α mediates vascular maintenance⁹⁷. Reduced HIF-1 α levels in mice impair embryonic vascular development, revascularization of ischaemic tissues, and angiogenesis in injured tissues and tumours⁹⁷. The use of HIF-1 α inhibitors to block tumour or ocular angiogenesis has therefore received attention. Conversely, *Hif1a* gene transfer in mice or activation of HIF-1 α by pharmacological blockade of PHDs promotes ischaemic tissue revascularization.

HIF-1 α also regulates tumour angiogenesis indirectly, by releasing chemoattractants such as SDF-1 α to recruit pro-angiogenic BMDCs⁶⁵. Gene silencing of *Phd2* in mouse tumour cells enhances vessel growth by similar mechanisms. Hypoxia also regulates the polarization and pro-angiogenic activity of tumour-associated macrophages (TAMs) by

means of HIF-1 α and HIF-2 α with different effects⁹⁶. That hypoxia and inflammation are closely intertwined is illustrated by the finding that signalling by HIF-1 α and nuclear factor- κ B cross-activate each other. In certain cases, hypoxic upregulation of VEGF occurs independently of HIF-1 α , and is mediated by the metabolic regulator PGC-1 α in preparation for oxidative metabolism once the ischaemic tissue is revascularized⁹⁸. Because HIF signalling contributes to acquired resistance against anti-VEGF therapy, the combined blockade of VEGF and HIF-1 α is being explored as a cancer treatment strategy.

There is increasing evidence for epigenetic control of angiogenesis, particularly by non-coding microRNAs (miRNAs)¹⁵, which induce messenger RNA degradation or block translation. Because miRNAs target multiple genes, they are well positioned to regulate complex processes such as angiogenesis. Endothelial cells express several miRNAs that are induced by hypoxia or VEGF. Most of those stimulate angiogenesis by hijacking pro-angiogenic cascades, while suppressing angiostatic pathways⁹⁹. The expression of miR-126 is induced by the mechanosensitive transcription factor KLF2A and integrates the mechanosensory stimulus of blood flow to shape the vascular system¹⁰⁰. Endothelial-cell-specific loss of DICER, an exonuclease involved in miRNA biogenesis, impairs pathological angiogenesis. Angiogenic miRNAs seem to offer significant pro- or anti-angiogenic potential.

plasmin also liberate angiogenic factors such as VEGF and FGF from immobilized matrix stores⁶³. VEGF isoforms that are cleaved by MMPs (and therefore soluble) preferentially enlarge vessels, whereas MMP-resistant matrix-bound VEGF supports vessel branching⁶⁴. Macrophages, neutrophils and mast cells initiate angiogenesis by MMP9-mediated activation of VEGF^{65,66}. Proteases such as MMP9 also participate in the mobilization of progenitors from the bone marrow by shedding soluble forms of membrane-bound cytokines (such as KIT ligand; also known as stem-cell factor or SCF)⁶⁷. MMPs establish a premetastatic niche by allowing the recruitment of marrow progenitors²⁹. Given their destructive potential, the activity of proteases must be tightly controlled. For instance, loss of the inhibitor PAI-1 prevents vessel branching because excessive ECM breakdown leaves no matrix support for the sprout⁶⁸. In addition, basement membrane deposition during vessel maturation requires the activity of MMP inhibitors such as TIMPs. Because degradation of ECM components can also generate anti-angiogenic fragments such as tumstatin and angiostatin¹, protease inhibitors must be judiciously evaluated for biological effects.

Junctional molecules

Cell–cell communication is fundamental for vessels to act as a synchronized unit along their longitudinal axis. Such coordination is accomplished by cell–cell communication through gap junctions, established by connexins, which inform upstream feeding vessels about the perfusion status of downstream tissues to prevent shunting, a well-known defect in tumour vessels⁶⁹. Apart from these long-range communication junctions, endothelial cells and pericytes have junctions for short-range communication.

Quiescent endothelial cells form a monolayer of interconnected cells, whereas angiogenic endothelial cells dissociate their junctions to migrate. The tight junctional molecules claudins, occludins and junctional adhesion molecules maintain barriers, such as the BBB, whereas adherens junctions establish cell–cell adhesion, cytoskeleton remodelling and intracellular signalling⁷⁰. Loss of VE-cadherin does not prevent vessel development, but induces defects in vascular remodelling and integrity². VE-cadherin is also required for localizing CD34 and its sialomucin receptor to cell–cell contacts for lumen formation⁷¹. In quiescent pericytes

endothelial cells, VE-cadherin promotes vessel stabilization by inhibiting VEGFR-2 signalling while activating TGF β pathways. Notably, oxygen sensors control VE-cadherin expression in a feedback loop, so vessel perfusion can be optimized when the oxygen supply is insufficient⁷². N-cadherin stabilizes contacts between endothelial cells and pericytes. During sprouting, the adhesive function of VE-cadherin between adjacent cells is reduced by endocytosis in response to VEGF and angiogenic factors⁷⁰. At the same time, the localization of VE-cadherin at filopodia allows tip cells to establish new contacts with cells on outreaching sprouts. Antibodies, recognizing neoepitopes of VE-cadherin that are exposed after dissociation of adherens junctions during sprouting, offer opportunities for selective blockage of endothelial cell growth without affecting endothelial cell maintenance.

Chemokines and G-protein-coupled receptors

Chemokines regulate angiogenesis by recruiting pro-angiogenic immune cells and endothelial progenitor cells, or through the direct activation of endothelial G-protein-coupled chemokine receptors (GPCRs). A well-known chemokine is SDF-1 α , which binds to its receptor CXCR4 on tip cells⁷³. SDF-1 α is upregulated by HIF-1 α in hypoxia, and supports mobilization and retention of pro-angiogenic CXCR4⁺ BMDCs to promote revascularization of ischaemic organs. Cancer-associated fibroblasts also release SDF-1 α . Another chemokine is the biologically active lipid sphingosine-1-phosphate (S1P), which binds to the S1P family of G-protein-coupled receptors (S1PRs) and regulates endothelial cell barrier function, vessel stability and angiogenesis, in part by crosstalking to PDGF and VEGF receptors — a process known as GPCR-jacking⁷⁴. Inhibitors of SDF-1 α , CXCR4 and S1P are being developed for cancer treatment⁷⁵. Another recently identified GPCR is GPR124, which regulates BBB differentiation⁷⁵.

Other pathways and challenges in translation

Other pathways also regulate angiogenesis, some of which provide guidance signals to navigating tip cells (Box 2). Given the ancestral function of vessels to supply oxygen, vessel formation is under the control of oxygen sensors (Box 3). As already mentioned, various anti-angiogenic avenues in addition to VEGF blockade are under development¹⁴. A challenge for

the future will be to identify optimal treatment regimens for these agents, either as monotherapy or in combination with VEGF blockade.

A major hurdle in translating the above-mentioned insights into clinically successful treatments stems from the fact that various anti-angiogenic approaches have different effects or are more effective in preclinical than clinical settings. This divergence may be due to several factors. First, most preclinical studies examine the effect of anti-angiogenic agents on transplantable, rapidly growing primary tumours, whereas most anti-angiogenic drugs have been approved for spontaneously arising, slowly evolving cancers in metastatic settings or for advanced disease in patients. The spontaneous tumour models in genetically engineered mice also do not recapitulate various aspects of the human disease. Differences in malignancy, vascularization and the stromal microenvironment between humans and mice lead to different responses. Second, only a few preclinical studies have analysed the effects of anti-angiogenic agents in residual disease after cytoreductive therapy or in the adjuvant setting, and often without chemotherapy. Third, the doses used in preclinical mouse studies are often higher than those given to patients, resulting in more pronounced antivasculature and antitumour effects in mice. Fourth, most genetic studies use mice in which the relevant angiogenic gene has been deleted before tumours become established, which is different from pharmacological intervention in patients after the cancer has become detectable. Finally, the dose and schedule of anti-angiogenic and chemotherapeutic drugs in the clinic have not been optimized, owing to cost and other considerations¹⁴.

Clinical anti-angiogenesis with VEGF blockers

Several VEGF blockers have been approved for clinical use in cancer and eye diseases^{6,7}. So far, the US Food and Drug Administration has approved the use of the VEGF-neutralizing antibody bevacizumab (Avastin) for metastatic colorectal cancer, metastatic non-squamous non-small-cell lung cancer, metastatic breast cancer, recurrent glioblastoma multiforme (GBM) and metastatic renal cell carcinoma (RCC) (Table 1). In addition, several multi-targeted TKIs, which block the signalling of pathways such as VEGF, have been approved, including sorafenib (Nexavar) for metastatic RCC and unresectable hepatocellular carcinoma, and sunitinib (Sutent) and pazopanib (Votrient) for metastatic RCC (Table 1). Recently, vandetanib (Zactima) has been approved for unresectable or metastatic medullary thyroid cancer and sunitinib has been recommended for approval for advanced pancreatic neuroendocrine tumours, but the clinical data have not yet been published. Treatment with VEGF inhibitors generally prolongs the survival of responsive patients with cancer of the order of months (Table 1). Two anti-VEGF compounds — intravitreal injection of the VEGF aptamer pegaptanib (Macugen) and the anti-VEGF Fab antibody ranibizumab (Lucentis) — have been approved for treatment of the wet (neovascular) form of age-related macular degeneration, which causes blindness owing to the formation of leaky neovessels. Bevacizumab is also used off-label for this condition.

Notwithstanding these successes, the clinical use of VEGF blockers in patients with cancer has shown that anti-angiogenic therapy is more challenging than anticipated. For example, VEGF receptor TKIs are effective as monotherapy in certain cancers, but fail in others or are toxic when combined with chemotherapy^{6,14}. The use of bevacizumab is approved only when combined with cytotoxic or cytokine therapy (with the exception of patients with GBM). Many patients with metastatic disease are refractory or acquire resistance to VEGF inhibitors⁴⁶, and biomarkers to identify responders are missing¹⁴. In a recent trial, bevacizumab prolonged disease-free progression but not overall survival in patients with metastatic RCC⁷⁶, and failed to show benefit in the adjuvant setting⁷⁷. Moreover, questions have begun to arise about whether anti-angiogenic therapy causes cancer cells to become more malignant^{78,79}. What are the reasons for these problems, and what can be done to move forward? The discussion in the next sections does not offer an answer to the daily challenges in oncological practice, but provides some avenues for developing future strategies.

Table 1 | Overview of anti-angiogenic drugs in cancer

Drug	Approved indication	Improvement in RR (%)	Improvement in PFS (months)	Improvement in OS (months)
Bevacizumab	Metastatic colorectal cancer (with chemotherapy)	10	4.4	4.7*
		0	1.4	1.4*
		7.8	2.8	2.5*
		14.1	2.6	2.1†
	Metastatic non-squamous NSCLC (with chemotherapy)	20	1.7	2.0*
		10.3–14.0	0.4–0.6	NR*
	Metastatic breast cancer (with chemotherapy)‡	15.7	5.9	NS*
		9–18	0.8–1.9	NS*
		11.8–13.4	1.2–2.9	NS*
		9.9	2.1	NS†
	Recurrent GBM (monotherapy)	Currently only phase II data reported		
Sunitinib	Metastatic RCC	18	4.8	NS*
		12.4	3.3	NS*
Sorafenib	Metastatic RCC	35	6.0	4.6*
	Unresectable HCC	8	2.7	NS†
		1	NS	2.8*
Pazopanib	Metastatic RCC	2	1.4	2.3*
		27	5.0	NR*†

Anti-angiogenic therapies currently approved by the US Food and Drug Administration (FDA) for treatment of malignancies. Per indication, the results of various trials are shown. The data show the improvement observed after the addition of the anti-VEGF therapy. GBM, glioblastoma multiforme; HCC, hepatocellular carcinoma; IFN, interferon; NR, not reported; NS, not significant; NSCLC, non-small-cell lung carcinoma; OS, overall survival; PFS, progression-free survival; RCC, renal cell carcinoma; RR, response rate. For reference, see <http://clinicaltrials.gov>.

*First-line therapy.

†Second-line therapy.

‡The FDA recommended the withdrawal of bevacizumab for breast cancer in December 2010; this is under appeal, with a hearing expected in June 2011. However, bevacizumab is approved for metastatic breast cancer in Europe, except in the United Kingdom.

Refractoriness to VEGF blockade in advanced cancer

A fraction of patients with cancer are refractory to VEGF-inhibitor treatment⁴⁶. The extent of refractoriness varies from one cancer to another, differs between micro- and macrometastatic disease, and differs for various types of VEGF blocker. Patients can be intrinsically refractory and never show any response to treatment, or develop evasive resistance during the course of treatment. Several mechanisms have been proposed to explain these phenomena, which are related to changes in the tumour cells, endothelial cells or other stromal cells^{6,14,46,80} (Fig. 3). It is important to note that these mechanisms have been identified for advanced, late-stage, macrometastatic disease only.

Tumour angiogenesis can become VEGF independent at a more advanced stage because of the production of other pro-angiogenic molecules, and thus respond poorly to VEGF blockade. Hypoxia induced by vessel regression after VEGF blockade can also switch on a more invasive and metastatic program, whereas in other cases, cancer (stem) cells can become hypoxia-tolerant when acquiring extra mutations and survive in poorly oxygenated niches. VEGF blockade inhibits sprouting angiogenesis, but may not be as efficient in suppressing other modes of tumour vascularization, relying on the recruitment of BMDCs, vessel co-option, vasculogenic mimicry or vessel splitting. Certain tumours, such as pancreatic carcinoma, contain a hypovascular stroma and are therefore less sensitive to anti-angiogenic agents. Vessel pruning by VEGF blockade can aggravate hypoxia, resulting in the upregulation of angiogenic factors such as PlGF, FGFs, chemokines and ephrins, and this may rescue tumour vascularization⁴⁶. Some tumour endothelial cells show signs of cytogenetic abnormalities and transforming stem-cell potential⁸¹, which could alter sensitivity to VEGF inhibition. Furthermore, GBM-like stem cells can differentiate into tumour endothelial cells, and VEGF blockers can only partially inhibit this process⁴.

Other stromal cells contribute to the resistance to VEGF blockade.

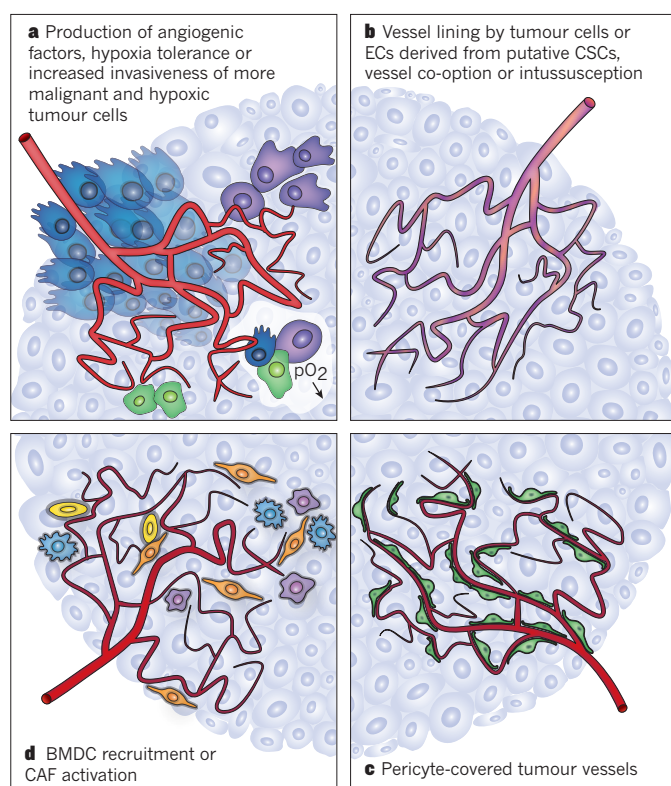


Figure 3 | Potential mechanisms of resistance to targeted VEGF therapy in cancer. Different mechanisms underlie the resistance to VEGF blockade seen in some patients with cancer. These mechanisms are not exclusive, and it is likely that several occur simultaneously in a single tumour. **a**, In established tumours, VEGF blockade aggravates hypoxia, which upregulates the production of other angiogenic factors or increases tumour cell invasiveness. Tumour cells that have acquired other mutations can also become hypoxia tolerant. The more malignant tumour cells are shown as dark green, blue and purple cells. **b**, Other modes of tumour vascularization, including intussusception, vasculogenic mimicry, differentiation of putative cancer stem cells (CSCs) into endothelial cells (ECs), vasculogenic vessel growth and vessel co-option (all denoted by the mosaic red–purple vessels), may be less sensitive to VEGF blockade. **c**, Tumour vessels covered by pericytes (green) are less sensitive to VEGF blockade. **d**, Recruited pro-angiogenic BMDCs (yellow), macrophages (blue and purple) or activated cancer-associated fibroblasts (CAFs; orange) can rescue tumour vascularization by the production of pro-angiogenic factors.

Hypoxia promotes the recruitment of angiocompetent BMDCs, including TEMs, TAMs, neutrophils, mast cells and CD11b⁺GR-1⁺ (also known as ITGAM⁺Ly6G⁺) myeloid-derived suppressor cells, which release angiogenic signals such as VEGF, BV8 (also known as PROK2) and MMPs⁸². VEGF blockade is often combined with chemotherapeutics — by sensitizing endothelial cells to cytotoxic damage, VEGF inhibitors impair endothelial cell survival and regrowth, but recruitment of BMDCs after chemotherapy can revascularize tumours ('vasculogenic rebound')^{6,83}. The release of angiogenic factors such as PDGF-CC by cancer-associated fibroblasts also contributes to resistance. Furthermore, vessels in most tumours are covered with few pericytes, but microvessels in some cancers acquire a dense pericyte coat with a thick basement membrane; such mature vessels are usually less sensitive to VEGF blockers^{31,32}. Understanding the molecular basis of these cancer-type-dependent resistance mechanisms against VEGF blockade offers opportunities to improve anti-angiogenic treatment.

VEGF blockers in the adjuvant setting

On the basis of the clinical experience with VEGF inhibitors in macrometastatic cancer, VEGF blockers were anticipated to be beneficial for micrometastatic disease in the adjuvant setting (that is, after surgical resection of the primary tumour). However, compared with

chemotherapy alone, adjuvant treatment of patients with micrometastatic disease in combination with bevacizumab and chemotherapy failed to prolong disease-free survival after three years⁷⁷. The administration of an anti-VEGF antibody initially prolonged disease-free survival in patients, but this benefit was lost after three years. The precise reasons for this remain unclear. It is possible that the micrometastatic tumour cells were less responsive to anti-VEGF treatment because they were in a state of angiogenic dormancy⁸⁴. The recruitment of pro-angiogenic BMDCs may convert micrometastasis to macrometastasis, but it unclear whether VEGF blockade eliminates this rescue pathway. Another hypothesis is that an angiogenesis rebound occurs after the arrest of anti-VEGF treatment, as documented in animal models⁸⁵. However, such vascular rebound did not occur after long-term treatment with a pan-VEGFR TKI in patients with GBM⁸⁶.

Another question is whether the transient disease-free survival benefit of anti-VEGF treatment was attributable to a change in the nature of the disease, and whether VEGF blockade caused the cancer to become more malignant after an initial delay. Some preclinical models show that VEGF blockade aggravates hypoxia and induces a pro-tumorigenic inflammatory state, which promotes invasiveness and metastasis, despite inhibition of primary tumour growth and prolongation of survival^{78,79}. However, another preclinical study reported no effect of VEGF blockade on metastasis in the adjuvant setting⁸⁷, and clinical trials have not shown an increase in malignancy or tumour-growth rebound after VEGF blockade, at least not in the metastatic setting. Moreover, a recent randomized phase II trial showed that continuous dosing and discontinuous dosing (four weeks on and two weeks off) of sunitinib have the same outcome in RCC patients. Finally, a meta-analysis of advanced cancers shows that VEGF blockade does not aggravate metastatic disease⁸⁸. Recurrent GBM is an exception, in which VEGF blockade increased tumour invasion, but even in these studies, tumours might have become more malignant because the treatment prolonged survival and allowed the cancer to progress further. Overall, there is an urgent need for an improved mechanistic understanding of vessel growth and resistance to anti-angiogenic therapy, particularly in micrometastatic lesions.

Tumour vessel abnormalities as a future target

Another parameter that could determine the overall efficiency of anti-VEGF therapy is the abnormal nature of tumour vessels. Tumour vessels become abnormal in almost all aspects of their structure and function⁸⁹. They are heterogeneous, tortuous, branch chaotically and have an uneven vessel lumen. In addition to abnormal endothelial cells, pericytes and the basement membrane are also abnormal. Owing to the leakiness of tumour vessels, escaping fluid raises the interstitial fluid pressure. As a result, blood flow is heterogeneous, and oxygen, nutrients, immune cells and drugs are distributed unevenly. Because radiation therapy and many chemotherapeutics rely on the formation of oxygen radicals to kill cancer cells, tumour hypoxia reduces their efficacy. These vessel abnormalities create a hostile milieu, characterized by hypoxia, low pH and high fluid pressure, which can select for more malignant cancer cells and lower barriers to their escape through leaky vessels.

These findings raise questions for the future. Excessive vessel pruning and growth arrest by anti-angiogenic agents could aggravate tumour invasiveness and metastasis by increasing hypoxia and creating a pro-tumorigenic inflammatory state. Vessel normalization could provide new therapeutic opportunities to slow down tumour invasiveness and dissemination, and increase tumour responses to chemotherapeutics and radiotherapy⁸⁹. Another consideration is how vessel normalization should be combined with an anti-angiogenic treatment. Vessel normalization was first recognized in mice xenografted with colon cancer and treated with an anti-VEGF antibody, but it is transient in mice and patients^{14,89,90}. Recent genetic studies in mice have shown that sustained vessel normalization can provide benefits. Indeed, haploinsufficiency of the oxygen-sensor PHD2 in endothelial cells induces sustained normalization of tumour vessels, without altering vessel density or size⁷². In these vessels, leakage, tortuosity and remodelling are reduced, whereas

endothelial cell quiescence, barrier tightening and vessel maturation are increased — changes that boost perfusion and decrease hypoxia⁷². A streamlined monolayer of phalanx endothelial cells is also formed, providing a more impenetrable barrier for intravasating tumour cells⁷². These changes do not affect tumour growth, but reduce tumour cell invasiveness, intravasation and metastasis⁷². Although these genetic studies offer an elegant example, the challenge will be to develop therapeutic strategies that translate these insights into daily practice in the clinic.

Future directions

An important question is how anti-angiogenic medicine can be improved. In the short term, the use of current anti-VEGF agents should be optimized. Given the low response rates, a step forward would be the discovery of predictive biomarkers to identify responders among the large patient group of non-responders. So far, only a few candidates for predictive biomarkers have been identified, but they emerged from small studies and require prospective validation in independent randomized trials¹⁴. Another consideration is the optimization of the dose and duration of anti-angiogenic drug delivery. Little is understood about the mechanisms of vascularization of micrometastatic lesions, and agents that can block other modes of tumour vascularization (such as co-option, intussusception, vasculogenesis and vasculogenic mimicry) are needed. Furthermore, understanding the mechanistic differences between VEGFR TKIs and anti-VEGF antibodies (for instance, whether the former are effective as monotherapy because they inhibit several targets, whereas the latter require combination chemotherapy in most instances) will help to optimize the design of anticancer treatments.

In the intermediate term, anti-VEGF agents could be combined with agents that target the escape pathways detected in clinical studies (not in mice). Examples are ANG-2, PlGF, SDF-1 α and CXCR4 (ref. 14). The challenge will be when to add these second agents — before, during or after anti-VEGF therapy. In the long term, the therapeutic potential of vascular normalization agents based on recently identified targets should be evaluated in preclinical models, but their clinical development will require years. By using combinatorial therapeutic approaches, it will be important to explore the eradication of most tumour vessels and normalization of residual vessels for longer durations than are now achievable with VEGF blockers alone. The potential of tumour vessel normalization to improve anticancer immune therapy should be explored further. Finally, it is important to test whether the approved anti-VEGF agents and those under development could be used to treat various non-malignant diseases characterized by abnormal vasculature, which afflict millions of people worldwide and in many cases have no effective treatment — such as age-related macular degeneration causing blindness, schwannomas causing loss of hearing, and atherosclerotic plaques causing stroke and myocardial infarction after rupture^{14,91}. A tight integration between preclinical and clinical research is crucial to achieve these goals. ■

1. Folkman, J. Angiogenesis: an organizing principle for drug discovery? *Nature Rev. Drug Discov.* **6**, 273–286 (2007).
2. Carmeliet, P. Angiogenesis in health and disease. *Nature Med.* **9**, 653–660 (2003).
3. Swift, M. R. & Weinstein, B. M. Arterial–venous specification during development. *Circ. Res.* **104**, 576–588 (2009).
4. Wang, R. *et al.* Glioblastoma stem-like cells give rise to tumour endothelium. *Nature* **468**, 829–833 (2010).
5. Schaper, W. Collateral circulation: past and present. *Basic Res. Cardiol.* **104**, 5–21 (2009).
6. Jain, R. K., Duda, D. G., Clark, J. W. & Loeffler, J. S. Lessons from phase III clinical trials on anti-VEGF therapy for cancer. *Nature Clin. Pract. Oncol.* **3**, 24–40 (2006).
7. Ferrara, N. VEGF-A: a critical regulator of blood vessel growth. *Eur. Cytokine Netw.* **20**, 158–163 (2009).
8. Nagy, J. A., Dvorak, A. M. & Dvorak, H. F. VEGF-A and the induction of pathological angiogenesis. *Annu. Rev. Pathol.* **2**, 251–275 (2007).
9. Neufeld, G. & Kessler, O. The semaphorins: versatile regulators of tumour progression and tumour angiogenesis. *Nature Rev. Cancer* **8**, 632–645 (2008).
10. Phng, L. K. & Gerhardt, H. Angiogenesis: a team effort coordinated by Notch. *Dev. Cell* **16**, 196–208 (2009).
11. Stockmann, C. *et al.* Deletion of vascular endothelial growth factor in myeloid cells accelerates tumorigenesis. *Nature* **456**, 814–818 (2008).
12. Lee, S. *et al.* Autocrine VEGF signaling is required for vascular homeostasis. *Cell* **130**, 691–703 (2007).

13. Lanahan, A. A. *et al.* VEGF receptor 2 endocytic trafficking regulates arterial morphogenesis. *Dev. Cell* **18**, 713–724 (2010).
 14. Jain, R. K. *et al.* Biomarkers of response and resistance to antiangiogenic therapy. *Nature Rev. Clin. Oncol.* **6**, 327–338 (2009).
- This review highlights the importance of various types of biomarker to personalize anti-angiogenic medicine, distinguish responders from non-responders and monitor treatment efficacy and adverse effects.**
15. Buysschaert, I., Schmidt, T., Roncal, C., Carmeliet, P. & Lambrechts, D. Genetics, epigenetics and pharmaco-(epi)genomics in angiogenesis. *J. Cell. Mol. Med.* **12**, 2533–2551 (2008).
 16. Tvorogov, D. *et al.* Effective suppression of vascular network formation by combination of antibodies blocking VEGFR ligand binding and receptor dimerization. *Cancer Cell* **18**, 630–640 (2010).
 17. Tammela, T. & Alitalo, K. Lymphangiogenesis: molecular mechanisms and future promise. *Cell* **140**, 460–476 (2010).
 18. Herbert, S. P. *et al.* Arterial–venous segregation by selective cell sprouting: an alternative mode of blood vessel formation. *Science* **326**, 294–298 (2009).
 19. Fischer, C., Mazzone, M., Jonckx, B. & Carmeliet, P. FLT1 and its ligands VEGFB and PlGF: drug targets for anti-angiogenic therapy? *Nature Rev. Cancer* **8**, 942–956 (2008).
 20. Carmeliet, P. *et al.* Synergism between vascular endothelial growth factor and placental growth factor contributes to angiogenesis and plasma extravasation in pathological conditions. *Nature Med.* **7**, 575–583 (2001).
 21. Rolny, C. *et al.* HRG inhibits tumor growth and metastasis by inducing macrophage polarization and vessel normalization through downregulation of PlGF. *Cancer Cell* **19**, 31–44 (2011).
 22. Van de Veire, S. *et al.* Further pharmacological and genetic evidence for the efficacy of PlGF inhibition in cancer and eye disease. *Cell* **141**, 178–190 (2010).
 23. Bais, C. *et al.* PlGF blockade does not inhibit angiogenesis during primary tumor growth. *Cell* **141**, 166–177 (2010).
 24. Hagberg, C. E. *et al.* Vascular endothelial growth factor B controls endothelial fatty acid uptake. *Nature* **464**, 917–921 (2010).
 25. Bry, M. *et al.* Vascular endothelial growth factor-B acts as a coronary growth factor in transgenic rats without inducing angiogenesis, vascular leak, or inflammation. *Circulation* **122**, 1725–1733 (2010).
 26. Schwartz, J. D., Rowinsky, E. K., Youssoufian, H., Pytowski, B. & Wu, Y. Vascular endothelial growth factor receptor-1 in human cancer. *Cancer* **116**, 1027–1032 (2010).
 27. Lichtenberger, B. M. *et al.* Autocrine VEGF signaling synergizes with EGFR in tumor cells to promote epithelial cancer development. *Cell* **140**, 268–279 (2010).
 28. Duda, D. G. & Jain, R. K. Premetastatic lung ‘niche’: is vascular endothelial growth factor receptor 1 activation required? *Cancer Res.* **70**, 5670–5673 (2010).
 29. Kaplan, R. N. *et al.* VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* **438**, 820–827 (2005).
 30. Jain, R. K. Molecular regulation of vessel maturation. *Nature Med.* **9**, 685–693 (2003).
 31. Hellberg, C., Ostman, A. & Heldin, C. H. PDGF and vessel maturation. *Recent Results Cancer Res.* **180**, 103–114 (2010).
 32. Gaengel, K., Genove, G., Armulik, A. & Betsholtz, C. Endothelial-mural cell signaling in vascular development and angiogenesis. *Arterioscler. Thromb. Vasc. Biol.* **29**, 630–638 (2009).
- This article discusses the role of pericytes in tumour angiogenesis, and highlights the possible pro-metastatic effects of blocking pericyte coverage.**
33. Quaegebeur, A., Segura, I. & Carmeliet, P. Pericytes: blood–brain barrier safeguards against neurodegeneration? *Neuron* **68**, 321–323 (2010).
 34. Song, S., Ewald, A. J., Stallcup, W., Werb, Z. & Bergers, G. PDGFR β perivascular progenitor cells in tumours regulate pericyte differentiation and vascular survival. *Nature Cell Biol.* **7**, 870–879 (2005).
 35. Bergers, G., Song, S., Meyer-Morse, N., Bergsland, E. & Hanahan, D. Benefits of targeting both pericytes and endothelial cells in the tumor vasculature with kinase inhibitors. *J. Clin. Invest.* **111**, 1287–1295 (2003).
 36. McCarty, M. F. *et al.* Overexpression of PDGF-BB decreases colorectal and pancreatic cancer growth by increasing tumor pericyte content. *J. Clin. Invest.* **117**, 2114–2122 (2007).
 37. Sennino, B. *et al.* Cellular source and amount of vascular endothelial growth factor and platelet-derived growth factor in tumors determine response to angiogenesis inhibitors. *Cancer Res.* **69**, 4527–4536 (2009).
 38. Nisancioglu, M. H., Betsholtz, C. & Genove, G. The absence of pericytes does not increase the sensitivity of tumor vasculature to vascular endothelial growth factor-A blockade. *Cancer Res.* **70**, 5109–5115 (2010).
 39. Gerhardt, H. & Semb, H. Pericytes: gatekeepers in tumour cell metastasis? *J. Mol. Med.* **86**, 135–144 (2008).
 40. Lebrin, F. *et al.* Thalidomide stimulates vessel maturation and reduces epistaxis in individuals with hereditary hemorrhagic telangiectasia. *Nature Med.* **16**, 420–428 (2010).
 41. Crawford, Y. *et al.* PDGF-C mediates the angiogenic and tumorigenic properties of fibroblasts associated with tumors refractory to anti-VEGF treatment. *Cancer Cell* **15**, 21–34 (2009).
 42. di Tomaso, E. *et al.* PDGF-C induces maturation of blood vessels in a model of glioblastoma and attenuates the response to anti-VEGF treatment. *PLoS ONE* **4**, e5123 (2009).
 43. Pardali, E., Goumans, M. J. & ten Dijke, P. Signaling by members of the TGF- β family in vascular morphogenesis and disease. *Trends Cell Biol.* **20**, 556–567 (2010).
 44. Beenken, A. & Mohammadi, M. The FGF family: biology, pathophysiology and therapy. *Nature Rev. Drug Discov.* **8**, 235–253 (2009).
 45. Murakami, M. *et al.* The FGF system has a key role in regulating vascular

- integrity. *J. Clin. Invest.* **118**, 3355–3366 (2008).
46. Bergers, G. & Hanahan, D. Modes of resistance to anti-angiogenic therapy. *Nature Rev. Cancer* **8**, 592–603 (2008).
47. Augustin, H. G., Koh, G. Y., Thurston, G. & Alitalo, K. Control of vascular morphogenesis and homeostasis through the angiopoietin–Tie system. *Nature Rev. Mol. Cell Biol.* **10**, 165–177 (2009).
48. Saharinen, P. *et al.* Angiopoietins assemble distinct Tie2 signalling complexes in endothelial cell–cell and cell–matrix contacts. *Nature Cell Biol.* **10**, 527–537 (2008).
49. De Palma, M. *et al.* Tie2 identifies a hematopoietic lineage of proangiogenic monocytes required for tumor vessel formation and a mesenchymal population of pericyte progenitors. *Cancer Cell* **8**, 211–226 (2005).
50. Falcon, B. L. *et al.* Contrasting actions of selective inhibitors of angiopoietin-1 and angiopoietin-2 on the normalization of tumor blood vessels. *Am. J. Pathol.* **175**, 2159–2170 (2009).
51. Koh, Y. J. *et al.* Double antiangiogenic protein, DAA, targeting VEGF-A and angiopoietins in tumor angiogenesis, metastasis, and vascular leakage. *Cancer Cell* **18**, 171–184 (2010).
52. Benedito, R. *et al.* The Notch ligands Dll4 and Jagged1 have opposing effects on angiogenesis. *Cell* **137**, 1124–1135 (2009).
53. Phng, L. K. *et al.* Nrarp coordinates endothelial Notch and Wnt signaling to control vessel density in angiogenesis. *Dev. Cell* **16**, 70–82 (2009).
54. Jakobsson, L. *et al.* Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nature Cell Biol.* **12**, 943–953 (2010).
55. Thurston, G., Noguera-Troise, I. & Yancopoulos, G. D. The Delta paradox: DLL4 blockade leads to more tumour vessels but less tumour growth. *Nature Rev. Cancer* **7**, 327–331 (2007).
- References 10 and 55 discuss the molecular model of vessel branching, and the role of DLL4 and NOTCH signalling in tip- and stalk-cell formation.**
56. Yan, M. *et al.* Chronic DLL4 blockade induces vascular neoplasms. *Nature* **463**, E6–E7 (2010).
57. Corada, M. *et al.* The Wnt/ β -catenin pathway modulates vascular remodeling and specification by upregulating Dll4/Notch signaling. *Dev. Cell* **18**, 938–949 (2010).
58. Dejana, E. The role of Wnt signaling in physiological and pathological angiogenesis. *Circ. Res.* **107**, 943–952 (2010).
59. Desgrosellier, J. S. & Chersesh, D. A. Integrins in cancer: biological implications and therapeutic opportunities. *Nature Rev. Cancer* **10**, 9–22 (2010).
60. Hodivala-Dilke, K. α v β 3 integrin and angiogenesis: a moody integrin in a changing environment. *Curr. Opin. Cell Biol.* **20**, 514–519 (2008).
61. Contois, L., Akalu, A. & Brooks, P. C. Integrins as ‘functional hubs’ in the regulation of pathological angiogenesis. *Semin. Cancer Biol.* **19**, 318–328 (2009).
62. Deryugina, E. I. & Quigley, J. P. Pleiotropic roles of matrix metalloproteinases in tumor angiogenesis: contrasting, overlapping and compensatory functions. *Biochim. Biophys. Acta* **1803**, 103–120 (2010).
63. Bergers, G. *et al.* Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nature Cell Biol.* **2**, 737–744 (2000).
64. Iruela-Arispe, M. L. & Davis, G. E. Cellular and molecular mechanisms of vascular lumen formation. *Dev. Cell* **16**, 222–231 (2009).
65. Du, R. *et al.* HIF1 α induces the recruitment of bone marrow-derived vascular modulatory cells to regulate tumor angiogenesis and invasion. *Cancer Cell* **13**, 206–220 (2008).
66. Heissig, B. *et al.* Role of neutrophil-derived matrix metalloproteinase-9 in tissue regeneration. *Histol. Histopathol.* **25**, 765–770 (2010).
67. Heissig, B. *et al.* Recruitment of stem and progenitor cells from the bone marrow niche requires MMP-9 mediated release of kit-ligand. *Cell* **109**, 625–637 (2002).
68. Blasi, F. & Carmeliet, P. uPAR: a versatile signalling orchestrator. *Nature Rev. Mol. Cell Biol.* **3**, 932–943 (2002).
69. Pries, A. R., Hopfner, M., le Noble, F., Dewhirst, M. W. & Secomb, T. W. The shunt problem: control of functional shunting in normal and tumour vasculature. *Nature Rev. Cancer* **10**, 587–593 (2010).
70. Dejana, E., Tournier-Lasserre, E. & Weinstein, B. M. The control of vascular integrity by endothelial cell junctions: molecular basis and pathological implications. *Dev. Cell* **16**, 209–221 (2009).
71. Strlic, B. *et al.* The molecular basis of vascular lumen formation in the developing mouse aorta. *Dev. Cell* **17**, 505–515 (2009).
72. Mazzone, M. *et al.* Heterozygous deficiency of *PHD2* restores tumor oxygenation and inhibits metastasis via endothelial normalization. *Cell* **136**, 839–851 (2009).
- This article provides genetic evidence of a role for the oxygen-sensor PHD2 in tumour vessel normalization through its regulation of endothelial cell phalanx formation.**
73. Duda, D. G. *et al.* CXCL12 (SDF1 α) — CXCR4/CXCR7 pathway inhibition: an emerging sensitizer for anti-cancer therapies? *Clin. Cancer Res.* **17**, 2074–2080 (2011).
74. Visentin, B. *et al.* Validation of an anti-sphingosine-1-phosphate antibody as a potential therapeutic in reducing growth, invasion, and angiogenesis in multiple tumor lineages. *Cancer Cell* **9**, 225–238 (2006).
75. Kuhnert, F. *et al.* Essential regulation of CNS angiogenesis by the orphan G protein-coupled receptor GPR124. *Science* **330**, 985–989 (2010).
76. Escudier, B. *et al.* Phase III trial of bevacizumab plus interferon α -2a in patients with metastatic renal cell carcinoma (AVOREN): final analysis of overall survival. *J. Clin. Oncol.* **28**, 2144–2150 (2010).
77. Allegra, C. *et al.* Phase III trial assessing bevacizumab in stages II and III carcinoma of the colon: results of NSABP protocol C-08. *J. Clin. Oncol.* **29**, 11–16 (2011).
78. Ebos, J. M. *et al.* Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. *Cancer Cell* **15**, 232–239 (2009).
79. Paez-Ribes, M. *et al.* Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. *Cancer Cell* **15**, 220–231 (2009).
- References 78 and 79 provide preclinical evidence that VEGF or VEGFR blockade can enhance metastasis, whereas references 87 and 88 provide preclinical and clinical evidence that VEGF or VEGFR blockade has no effect on metastasis.**
80. Carmeliet, P. Angiogenesis in life, disease and medicine. *Nature* **438**, 932–936 (2005).
81. Hida, K. *et al.* Tumor-associated endothelial cells with cytogenetic abnormalities. *Cancer Res.* **64**, 8249–8255 (2004).
82. Ferrara, N. Pathways mediating VEGF-independent tumor angiogenesis. *Cytokine Growth Factor Rev.* **21**, 21–26 (2010).
- References 46 and 82 discuss the mechanisms of resistance to VEGF or VEGFR blockade.**
83. Nagengast, W. B. *et al.* VEGF-PET imaging is a noninvasive biomarker showing differential changes in the tumor during sunitinib treatment. *Cancer Res.* **71**, 143–153 (2011).
84. Joyce, J. A. & Pollard, J. W. Microenvironmental regulation of metastasis. *Nature Rev. Cancer* **9**, 239–252 (2009).
85. Mancuso, M. R. *et al.* Rapid vascular regrowth in tumors after reversal of VEGF inhibition. *J. Clin. Invest.* **116**, 2610–2621 (2006).
86. di Tomaso, E. *et al.* Glioblastoma recurrence after cediranib therapy in patients: lack of “rebound” revascularization as mode of escape. *Cancer Res.* **71**, 19–28 (2011).
87. Padera, T. P. *et al.* Differential response of primary tumor versus lymphatic metastasis to VEGFR-2 and VEGFR-3 kinase inhibitors cediranib and vandetanib. *Mol. Cancer Ther.* **7**, 2272–2279 (2008).
88. Miles, D. *et al.* Disease course patterns after discontinuation of bevacizumab: pooled analysis of randomized phase III trials. *J. Clin. Oncol.* **29**, 83–88 (2010).
89. Jain, R. K. Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy. *Science* **307**, 58–62 (2005).
- This review presents evidence for vascular abnormalities in tumours, and the therapeutic potential to normalize these vessels for treatment of cancer and other diseases characterized by abnormal vessels.**
90. Jain, R. K. Normalizing tumor vasculature with anti-angiogenic therapy: a new paradigm for combination therapy. *Nature Med.* **7**, 987–989 (2001).
91. Plotkin, S. R. *et al.* Hearing improvement after bevacizumab in patients with neurofibromatosis type 2. *N. Engl. J. Med.* **361**, 358–367 (2009).
92. Butler, J. M., Kobayashi, H. & Rafii, S. Instructive role of the vascular niche in promoting tumour growth and tissue repair by angiocrine factors. *Nature Rev. Cancer* **10**, 138–146 (2010).
- This review discusses the emerging concept that endothelial cells produce instructive signals for organ development, repair and even cancer. It also highlights the importance of hypoxia-mediated signalling in angiogenesis.**
93. Adams, R. H. & Eichmann, A. Axon guidance molecules in vascular patterning. *Cold Spring Harb. Perspect. Biol.* **2**, a001875 (2010).
94. Mosch, B., Reissenweber, B., Neuber, C. & Pietzsch, J. Eph receptors and ephrin ligands: important players in angiogenesis and tumor angiogenesis. *J. Oncol.* **2010**, 135285 (2010).
95. Sawamiphak, S. *et al.* Ephrin-B2 regulates VEGFR2 function in developmental and tumour angiogenesis. *Nature* **465**, 487–491 (2010).
96. Majumdar, A. J., Wong, W. J. & Simon, M. C. Hypoxia-inducible factors and the response to hypoxic stress. *Mol. Cell* **40**, 294–309 (2010).
97. Fraisl, P., Mazzone, M., Schmidt, T. & Carmeliet, P. Regulation of angiogenesis by oxygen and metabolism. *Dev. Cell* **16**, 167–179 (2009).
98. Arany, Z. *et al.* HIF-independent regulation of VEGF and angiogenesis by the transcriptional coactivator PGC-1 α . *Nature* **451**, 1008–1112 (2008).
- References 24, 97 and 98 discuss the link between metabolism and angiogenesis, and show that metabolic regulators control angiogenesis.**
99. Ohtani, K. & Dimmeler, S. Control of cardiovascular differentiation by microRNAs. *Basic Res. Cardiol.* **106**, 5–11 (2010).
- This review discusses the emerging evidence for regulation of angiogenesis by epigenetic mechanisms.**
100. Nicoli, S. *et al.* MicroRNA-mediated integration of haemodynamics and Vegf signalling during angiogenesis. *Nature* **464**, 1196–1200 (2010).

Acknowledgements We would like to acknowledge the late J. Folkman for inspiring us and others to work in the area of tumour angiogenesis. We would like to thank L. Claesson-Welch, H. Augustin, E. Dejana, S. Kozin, D. G. Duda, S. Goel, L. L. Munn, G. Sledge, R. Stupp and H. D. Suit for their comments, and L. Notebaert for help with the illustrations. We apologize to the authors whose work we could not cite because of the limit on the number of references. The work of P.C. is supported by a Federal Government Belgium grant (IUAP06/30), long-term structural Methusalem funding by the Flemish Government, a Concerted Research Activities Belgium grant (GOA2006/11), Leducq Transatlantic Network ARTEMIS and a grant from Flanders Research Foundation (FWO G.0673.08). The research of R.K.J. is supported by US National Institutes of Health grants P01-CA80124, R01-CA85140, R01-CA115767 and R01-CA126642, Federal Share/NCI Proton Beam Program Income, the National Foundation for Cancer Research and a Department of Defense Breast Cancer Research Innovator Award (W81XWH-10-1-0016).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to P.C. (peter.carmeliet@vib-kuleuven.be) or R.K.J. (jain@steele.mgh.harvard.edu).

Lessons on the pathogenesis of aneurysm from heritable conditions

Mark E. Lindsay^{1,2} & Harry C. Dietz^{2,3}

Aortic aneurysm is common, accounting for 1–2% of all deaths in industrialized countries. Early theories of the causes of human aneurysm mostly focused on inherited or acquired defects in components of the extracellular matrix in the aorta. Although several mutations in the genes encoding extracellular matrix proteins have been recognized, more recent discoveries have shown important perturbations in cytokine signalling cascades and intracellular components of the smooth muscle contractile apparatus. The modelling of single-gene heritable aneurysm disorders in mice has shown unexpected involvement of the transforming growth factor- β cytokine pathway in aortic aneurysm, highlighting the potential for new therapeutic strategies.

Aneurysm or arterial enlargement is the gross phenotype that manifests progressive organ failure of large arteries, including the aorta. Aortic aneurysm is usually not inherently dangerous, but enlarged arteries show a predisposition for tear (known as dissection), with high mortality rates. Clinicians recognize two predominant spatial distributions of aortic aneurysms in patient groups. The most common form, abdominal aortic aneurysm (AAA), is typically associated with advanced age and atherosclerosis, with attendant risk factors such as hypercholesterolaemia, hypertension and/or diabetes. These lesions are pathologically characterized by atheromata, invasion of inflammatory cells, destructive extracellular matrix remodelling, and depletion and dysfunction of vascular smooth muscle cells (VSMCs). Although it is clear that genetic determinants influence the development of AAA, there has been no description of a single major gene or locus effect that is sufficient to cause isolated abdominal aneurysm (that is, in the absence of evidence for a more systemic vasculopathy). The weight of evidence therefore indicates that AAA is a complex disorder that integrates the influence of predisposing genes with lifestyle-associated risk factors — analogous to coronary artery disease.

The second common site for aneurysm is the thoracic ascending aorta. Unlike AAA, thoracic aortic aneurysm (TAA) occurs in all age groups, is more highly associated with hereditary influences and does not show obligate association with cardiovascular risk factors. Pathologically, inherited forms of TAA typically show destructive matrix remodelling with elastin fragmentation, proliferation of VSMCs and a less prominent inflammatory component without atheromata. Many presentations of TAA show classic Mendelian inheritance with high or complete penetrance, suggesting the major contribution of a single gene. Familial TAA can be subdivided into syndromic presentations that show prominent features of a systemic connective tissue disorder (such as Marfan syndrome (MFS) and Loeys–Dietz syndrome (LDS)) and non-syndromic presentations (such as bicommissural aortic valve with TAA, and isolated familial TAA). Both the syndromic and non-syndromic groups include many disorders in which disease predominates in the very proximal ascending thoracic aorta (Fig. 1). In fact, there are only rare exceptions in which the ascending aorta is infrequently involved — such as vascular Ehlers–Danlos syndrome. Table 1 includes a list of human and mouse single-gene disorders associated with aneurysm predisposition.

Over the past couple of decades, the genes responsible for heritable aortic aneurysm have been identified at an accelerating pace. Gene identification has allowed the creation of mouse models of inherited

aortic aneurysm, providing the first opportunity to temporally and comprehensively interrogate the pathogenic sequence of aneurysm, extending from predisposition to clinical consequence, in an experimental context that mimics the physiological complexity of the human system. This combination of human molecular genetics and animal modelling has shown the involvement of diverse cytokine pathways, prominently the role of the transforming growth factor- β (TGF- β) pathway in aortic aneurysm. This Review focuses on the mechanisms of failure of large vessel wall homeostasis that challenge or inform historical perspectives, attempts to integrate emerging models of disease, and discusses the remaining challenges and opportunities in aneurysm research.

Excessive focus on elastin in TAA pathogenesis

There has been disproportionate historical focus on elastic fibres in pathogenetic models of inherited TAA. This derives from the near-uniform histological observation of reduced elastin content and elastic fibre fragmentation in the aortic media (the middle aortic layer), known as cystic medial necrosis. However, many elastin-deficiency states do not associate with aneurysm as a prominent phenotype. Although aortic aneurysm is an extremely rare manifestation of cutis laxa syndromes caused by mutations in the elastin gene¹, it is not observed in mice or humans with dominant and recessive forms of cutis laxa caused by deficiency of fibulin-5 (ref. 2), a crucial mediator of elastogenesis. By contrast, aneurysmal disease, including prominent involvement of the ascending aorta, is highly penetrant in inherited cutis laxa caused by fibulin-4 deficiency³. Humans and mice with a fibrillin-1 deficiency show failed elastic fibre homeostasis and highly penetrant aortic root aneurysms in the context of MFS (see ref. 4 and references therein). Pseudoxanthoma elasticum caused by *ABCC6* deficiency in humans also shows postnatally acquired elastic fibre fragmentation in the aorta, but does not typically show aneurysm⁵. These observations raise the important question of which molecular events, besides elastin-related issues, are common to the fibrillin-1- and fibulin-4-deficiency states but not recapitulated by fibulin-5 deficiency. The focus on the aortic media and elastic fibres may have been a distraction, and further insight could come from considering other distinguishing features of the proximal ascending aorta, including developmental cellular ontology. Genes mutated in MFS and vascular Ehlers–Danlos syndrome (*FBN1* and *COL3A1*, respectively) encode extracellular matrix elements (fibrillin-1 and collagen α -1(III), respectively). This discovery led to the generation of pathogenetic models that singularly invoke inherent structural weakness of the tissues. Such a view

¹Division of Pediatric Cardiology, Department of Pediatrics, Johns Hopkins Medical Institutions, Baltimore, Maryland 21205-1832, USA. ²McKusick–Nathans Institute of Genetic Medicine, Johns Hopkins Medical Institutions, Baltimore, Maryland 21205-1832, USA. ³Howard Hughes Medical Institute, Baltimore, Maryland, 21205-1832, USA.

boded poorly for the development of productive medical treatment strategies, requiring a means to alter the structural composition of inherently weak tissues. Fortunately, for both patients and researchers, the story of hereditary aneurysm has turned out to be much more complex and potentially permissive for therapeutic intervention.

MFS provides a link between aneurysm and TGF- β

A shift in thinking about the pathogenesis of aortic aneurysm occurred during the study of MFS⁴. Although many of the clinical manifestations of MFS could be caused by simple tissue weakness imposed by fibrillin-1 deficiency (such as aortic aneurysm, eye lens dislocation and emphysema), others were not so easily reconciled (bone overgrowth, craniofacial alterations and myxomatous valve disease). Insight came from studying lung disease in fibrillin-1-deficient mice. Contrary to expectation, mouse models of MFS did not show destructive and inflammatory emphysema. Instead, they showed primary failure of distal alveolar septation during late embryogenesis and the perinatal period⁶.

Mechanistic hypotheses built on the observation that fibrillin proteins show marked homology to latent TGF- β -binding proteins (LTBPs). Most TGF- β is secreted from cells in the context of a large latent complex (LLC) that includes the mature cytokine, a dimer of its processed amino-terminal propeptide (latency-associated peptide) and one of three LTBP isoforms (LTBP-1, LTBP-3 or LTBP-4). LTBPs target the TGF- β LLC to binding partners such as fibronectin and microfibrils composed of fibrillin-1 — an event that is thought to regulate TGF- β bioavailability and activity by controlling access to, or the efficiency of, TGF- β activators. One hypothesis was that failed or improper matrix sequestration of the LLC owing to fibrillin-1 deficiency could lead to promiscuous TGF- β activation. In keeping with this concept, the developing lungs of fibrillin-1-deficient mice showed decreased LLC levels but raised levels of free TGF- β and increased TGF- β signalling, as demonstrated by the nuclear translocation and phosphorylation of receptor-activated SMAD proteins 2 and 3 (pSMAD2/3). Notably, distal alveolar septation could be restored in fibrillin-1-deficient mice by the administration of a pan-specific anti-TGF- β neutralizing antibody⁶. Other manifestations of MFS, including myxomatous mitral valve disease, skeletal myopathy and, most importantly, aortic root aneurysm, were associated with increased TGF- β signalling in mouse models of MFS, and were attenuated or prevented by TGF- β antagonism with a neutralizing anti-TGF- β antibody in these mice *in vivo*^{4,7}.

TGF- β receptor mutations

Perhaps the most direct evidence of a major role for TGF- β in aneurysm pathogenesis came from the finding that mutations in the *TGFBR1* and *TGFBR2* genes — which encode the TGF- β receptor subunits TGFR-1 (also known as ALK-5) and TGFR-2, respectively — result in aneurysm conditions that have undeniable phenotypic overlap with MFS, a notable example of which is LDS^{8,9} (Table 1). Similar to MFS, patients with LDS typically show skeletal involvement, including long fingers, chest wall deformity and scoliosis. Other shared features include widening of the dural sac, skin stretch marks and mitral valve prolapse. Patients with LDS typically do not show lens dislocation, and can show many discriminating systemic features such as widely spaced eyes (hypertelorism), cleft palate or bifid uvula, cervical spine malformation or instability, osteoporosis and club foot deformity^{8,10}. Most importantly, patients with LDS show highly penetrant arterial tortuosity (an elongation of an artery resulting in a twisted course) and a strong predisposition for aneurysm and dissection throughout the arterial tree. Vascular disease in patients with LDS is more aggressive than in those with MFS, with rupture at a younger age (as young as 6 months) and at smaller aortic dimensions¹⁰.

Nearly all patients with LDS are heterozygous for missense substitutions in the kinase domain of TGFR-1 or TGFR-2. Recombinant expression of mutant receptors in cells naive for the corresponding receptor subunit failed to support TGF- β signalling, leading to the hypothesis that haploinsufficiency was the relevant mechanism⁹.

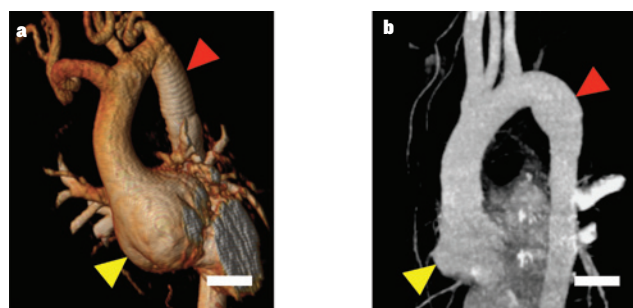


Figure 1 | Sites of TAA in transforming growth factor- β vasculopathy syndromes. Multidetector computer tomographic reconstruction images are shown. **a**, The thoracic aorta of a 9-year-old patient with LDS after emergent repair of a proximal descending (Stanford type B) aortic dissection, with a large unrepaired aortic root aneurysm (yellow arrowhead). The red arrowhead denotes the dacron tube graft. **b**, The thoracic aorta of a 16-year-old patient with LDS after surgical repair of an aortic root aneurysm (yellow arrowhead), with a discrete fusiform aneurysm of the proximal descending aorta (red arrowhead). Scale bars, 3 cm.

If this were the case, however, the identification of early nonsense mutations that elicit messenger RNA clearance through nonsense-mediated mRNA decay or whole allele deletions might be expected. Instead, the skewed mutational repertoire seems to manifest selection for receptor variants that traffic to the cell surface and bind ligand but lack the ability to propagate signal. Co-transfection studies in human cells using equimolar concentrations of normal and mutant receptors and experiments with heterozygous patient cells have shown apparent preservation of the signalling potential of wild-type receptor subunits, excluding a conventional dominant-negative mechanism^{8,9}. Notably, patient vascular tissue obtained at surgery or autopsy has consistently shown paradoxically enhanced TGF- β signalling, demonstrated by nuclear accumulation of pSMAD2 in VSMCs and increased output of TGF- β -driven gene products such as collagens and connective tissue growth factor (CTGF)^{8,10}. The architectural changes seen in the aortic wall of patients with LDS are highly reminiscent of those seen in MFS and other inherited forms of aortic root aneurysm. The mechanisms underlying this paradoxical effect remain unknown, but potentially include altered receptor trafficking, impaired autoregulation of TGF- β signalling, alternative signalling cascades or non-autonomous cellular events (see ‘Cancer biology as a guide to aneurysm’). Heterozygous loss-of-function *SMAD3* mutations recapitulate both the LDS phenotype and paradoxical enhancement of TGF- β signalling in the aortic wall¹¹.

Other TGF- β links to aneurysm

Several other aneurysmal disorders have been linked to TGF- β signalling⁴. High levels of TGF- β signalling, as assessed by nuclear pSMAD2 accumulation, have been observed in surgical samples from patients with diverse aneurysm conditions such as isolated familial TAA and bicommissural aortic valve with TAA⁴. Patients with autosomal recessive arterial tortuosity syndrome show diffuse and severe arterial tortuosity that is often associated with vascular stenoses, segmental vascular hypoplasia and arterial aneurysms, caused by loss-of-function mutations in the *SLC2A10* (also known as *GLUT10*) gene, which encodes the integral membrane protein glucose transporter type 10 (ref. 12). Vascular tissue from patients with arterial tortuosity syndrome shows the same signature of high TGF- β signalling seen in tissue from patients with MFS or LDS. Although the mechanism is poorly understood, cultured fibroblasts from patients with arterial tortuosity syndrome show impaired expression of the proteoglycan decorin — an antagonist of TGF- β superfamily signalling¹². Deficiency of the extracellular protein fibulin-4 causes autosomal recessive cutis laxa in association with arterial tortuosity and aortic aneurysm in both humans and mice^{3,13}. Curiously, fibulin-5 deficiency causes cutis laxa with arterial tortuosity but not aneurysm. Both fibulin-4 and fibulin-5 bind to fibrillin-1 and elastin, and deficiency in fibulin-4 or

Table 1 | Human hereditary aneurysm conditions and mouse models of aneurysm

Gene (protein)	Human aneurysmal syndrome	Animal-model phenotype	Pathway
Extracellular matrix protein			
<i>FBN1</i> (fibrillin-1)	MFS; highly penetrant ascending aortic aneurysm	KO: perinatal lethal, pulmonary hypoplasia, arteriopathy; hypomorphic: arteriopathy, aneurysm, dissection and systemic manifestations of MFS	TGF- β ^{54,55}
<i>EFEMP2</i> (fibulin-4)	Cutis laxa with aneurysm; ascending aortic aneurysm and tortuosity	KO: ascending aortic aneurysm, defective elastogenesis, perinatal lethality	TGF- β ^{3,13}
<i>ELN</i> (elastin)	Cutis laxa with aneurysm; low penetrance ascending aortic aneurysm and dissection	Haploinsufficient: obstructive arterial disease with increased VSMC proliferation, increased lamellae number; KO: accentuated phenotype	Unknown ^{1,56}
<i>COL1A1</i> (collagen α -1(I))	Osteogenesis imperfecta; extremely rare aortic aneurysm; EDS, type 7A; dissection of medium-sized arteries	KO: adult-onset aortic aneurysm and dissection	Collagen metabolism ^{57,58}
<i>COL1A2</i> (collagen α -2(I))	Osteogenesis imperfecta; extremely rare aortic aneurysm; EDS, cardiac valvular dystrophy type 7B; borderline aortic root enlargement with aortic regurgitation	Homozygous LOF: decreased body weight, bony abnormalities, no arterial phenotype reported	Collagen metabolism ⁵⁹
<i>COL3A1</i> (collagen α -1(III))	EDS, type 4; frequent arterial dissection with infrequent aneurysm	KO: frequent neonatal mortality, aortic rupture, intestinal rupture	Collagen metabolism ^{60,61}
<i>COL4A1</i> (collagen α -1(IV))	Hereditary angiopathy, nephropathy, aneurysms and muscle cramps; infrequent aneurysms	KO: embryonic lethal (E10.5–11.5), basement membrane failure	Collagen metabolism ^{62,63}
<i>COL4A5</i> (collagen α -5(IV))	X-linked Alport syndrome; ascending aortic and abdominal aneurysms and dissections	Nonsense mutation: no overt aortic disease noted	Collagen metabolism ⁶⁴
<i>LOX</i> (lysyl oxidase)	No human phenotype described	KO: low penetrance aortic aneurysm, perinatal lethality	Collagen metabolism; TGF- β ¹⁵
<i>PLOD1</i> (lysyl hydroxylase 1)	EDS, type 6; rare aneurysm	KO: spontaneous aneurysm and dissection, gait abnormalities	Collagen metabolism ^{65,66}
<i>PLOD3</i> (lysyl hydroxylase 3)	Bone fragility with contractures, arterial rupture and deafness; frequent medium-sized arterial aneurysms	KO: embryonic lethal (E9.5) and basement membrane fracture	Collagen metabolism ^{67,68}
Transmembrane protein			
<i>TGFBR1</i> (TGF- β receptor type 1)	LDS; highly penetrant root and diffuse large and medium arterial aneurysms	KO: midgestational death with yolk sac defects; M318R heterozygous knock-in: aortic root and diffuse aneurysm (D. Loch, unpublished observations)	TGF- β ^{8,69}
<i>TGFBR2</i> (TGF- β receptor type 2)	LDS; highly penetrant root and diffuse large and medium arterial aneurysms; familial thoracic aortic aneurysms and dissections; highly penetrant root and medium arterial aneurysms	KO: defects in haematopoiesis and vasculogenesis, embryonic lethal (E10.5); <i>Tgfb2</i> ^{lox} ; impaired elastogenesis, decreased lysyl oxidase in aorta; G357W heterozygous knock-in: aortic root and diffuse aneurysm (D. Loch, unpublished observations)	TGF- β ^{8,19,70}
<i>ENG</i> (endoglin)	Hereditary haemorrhagic telangiectasia; incompletely penetrant aortic and medium-sized arterial aneurysms	KO: defective vasculogenesis, embryonic lethal (E10); haploinsufficient: haemorrhagic telangiectasia causing strokes, fatal haemorrhage and heart failure	TGF- β superfamily ^{71,72}
<i>ACVRL1</i> (activin receptor-like kinase I)	Hereditary haemorrhagic telangiectasia; incompletely penetrant aortic and medium-sized arterial aneurysms	KO: defective vasculogenesis, embryonic lethal, excessive fusion of capillary plexuses; haploinsufficient: haemorrhagic telangiectasia	TGF- β superfamily ^{73,74}
<i>SLC2A10</i> (glucose transporter type 10)	Arterial tortuosity syndrome; diffuse arterial tortuosity, stenoses, aneurysms	Homozygote missense: arterial thickening with increased elastin deposition, elastin fractures at advanced age	TGF- β ^{12,75}
<i>NOTCH1</i> (NOTCH1)	Bicuspid valve with ascending aortic aneurysm	KO: embryonic lethal (E9.5), required for somite segmentation, defects in angiogenesis	NOTCH1–JAGGED1 ^{76,77}
<i>JAG1</i> (JAGGED1)	Alagille syndrome; intracranial aneurysms, coarctation of the aorta, aortic aneurysm	KO: embryonic lethal (E9.5) with diffuse haemorrhages	NOTCH1–JAGGED1 ^{78,79}
<i>GJA1</i> (connexin-43)	Hypoplastic left heart syndrome (HLHS)	Nonsense mutation (W45X): coronary artery aneurysms	Unknown ⁸³

fibulin-5 is associated with profound failure of elastogenesis — the probable cause of arterial tortuosity. Fibulin-5 mainly promotes elastin fibre assembly by the recruitment of tropoelastin to microfibrils¹⁴. By contrast, fibulin-4 is needed for the recruitment of lysyl oxidase (LOX), a copper-dependent enzyme that catalyses crosslinking of elastin molecules¹⁴. In keeping with these findings, LOX-deficient mice show severely disrupted aortic laminae, arterial tortuosity and low penetrance aneurysm¹⁵. Mice and humans deficient in fibulin-4 show increased TGF- β signalling in the vessel wall, which may be directly related to the ability of LOX to inhibit TGF- β enzymatically¹⁶. In this light, it seems that the aneurysm phenotype specific to fibulin-4 deficiency manifests a loss of a function other than elastin crosslinking, plausibly including TGF- β repression.

Despite several lines of evidence invoking high TGF- β signalling in aneurysm, conflicting observations exist. For example, high vascular TGF- β signalling was shown in *Emilin1*-deficient mice in association

with a diffusely small vascular system and hypertension in juvenile animals¹⁷. The developmental timing and vascular distribution of high TGF- β signalling were not assessed, and there was no comprehensive analysis for aneurysms¹⁷. Another study has shown that lineage-specific ablation of TGF- β signalling in VSMCs in the ascending aorta results in perturbations of vascular morphogenesis in fetal mice, including persistent truncus arteriosus, impaired elastogenesis and apparent vessel widening that was equated with aneurysm¹⁸; similar impairment of elastogenesis was seen only in the descending aorta of mice with global *Tgfb2* deletion in VSMCs¹⁹. There is further indirect evidence that low TGF- β signalling may also be involved in developmental presentations of aneurysm in MFS. An emerging view is that the fibrillin proteins have a dichotomous role in TGF- β regulation. Fibrillin-2 was shown to concentrate TGF- β ligands (prominently BMP7), and this is required to support morphogenetic events at sites of intended function. In the

Gene (protein)	Human aneurysmal syndrome	Animal-model phenotype	Pathway
Transmembrane protein cont.			
<i>PKD1</i> (polycystin-1)	Polycystic kidney disease with intracranial aneurysms	KO: embryonic lethal (E14.5) with polycystic kidneys; hypomorphic expression: adult-onset aortic aneurysm and dissection	mTOR ^{80,81}
<i>PKD2</i> (polycystin-2)	Polycystic kidney disease with intracranial aneurysms	KO: defects in cardiac septation and left–right axis determination, kidney and pancreatic cysts	Unknown ⁸²
Cytoplasmic protein			
<i>SMAD3</i> (SMAD family member 3)	LDS; aortic aneurysm with osteoarthritis	KO: metastatic colorectal cancer	TGF- β ¹¹
<i>ACTA2</i> (α -smooth muscle actin)	Familial aortic aneurysm with livedo reticularis and iris flocculi	KO: viable offspring with normal lifespan and impaired vascular contractility	IGF-1, Ang II ^{38,84}
<i>MYH11</i> (smooth muscle myosin)	Familial aortic aneurysm with patent ductus arteriosus	KO: neonatal lethality, urinary retention, dilated cardiomyopathy	IGF-1, Ang II ^{36,37,85}
<i>FLNA</i> (filamin-A)	Periventricular nodular heterotopia with EDS features; ascending aortic aneurysm and valvular dystrophy	KO: neonatal lethality, persistent truncus arteriosus, endothelial cell–cell contact defects	Unknown ^{40,86}
<i>NF1</i> (neurofibromin-1)	Neurofibromatosis; medium-sized arterial aneurysm and stenosis	KO: enlarged head, pale liver, cardiac malformations	Ras–MEK–ERK ³⁰
<i>PTPN11</i> (protein-tyrosine phosphatase 2C)	Noonan and LEOPARD syndromes; coronary artery aneurysms and rare ascending aortic aneurysm	KO: embryos die at preimplantation; missense mutation (D61G): cardiac defects, defective valvulogenesis, skeletal anomalies, myeloproliferative disorder	Ras–MEK–ERK ^{29,87,88}
<i>NPHP3</i> (nephrocystin-3)	Nephronophthisis	KO: low penetrance intracranial aneurysms	Unknown ⁸⁹
<i>NOS3</i> (nitric oxide synthase 3)	Refractory hypertension	KO: abnormal aortic development with bicuspid aortic valve; in combination with <i>Apoe</i> ^{−/−} , mice show abdominal arterial aneurysm and dissections	Nitric oxide ^{90,91}
<i>TSC2</i> (tuberin)	Tuberous sclerosis; diffuse thoracoabdominal aneurysms	Heterozygous KO: increased proliferation of VSMCs after injury	mTOR ⁹²
<i>GAA</i> (lysosomal α -glucosidase)	Acid maltase deficiency, adult onset; intracranial aneurysms	KO: lysosomal accumulation in heart, aorta, skeletal muscle	Unknown ⁹³
<i>S100A12</i> (S100A12)	No human phenotype; increased S100A12 protein expression in human <i>MYH11</i> -mutation aneurysmal tissues	Sm22 α promoter–S100A12 transgenic mouse: vascular smooth muscle disarray, elastin fragmentation, thoracic aneurysm	IL-6, TGF- β ³⁹
Nuclear protein			
<i>MED12</i> (mediator complex subunit 12)	Lujan–Fryns syndrome; extremely rare aneurysm	Hypomorphic mutants: embryonic lethal (E10), defects in neural tube closure, somatogenesis, heart formation	WNT– β -catenin, WNT–PCP ⁹⁴
<i>KLF15</i> (Krüppel-like factor 15)	No human phenotype; Krüppel-like factor 15 downregulated in human abdominal aortic aneurysm	KO: aortic aneurysm and cardiomyopathy	TSP-1, p53, TGF- β ⁹⁵
<i>KLF2</i> (Krüppel-like factor 2)	No human phenotype	KO: embryonic aortic aneurysm and dissection	Unknown ⁹⁶
Chromosomal anomaly			
45, X	Turner syndrome; bicuspid aortic valve, coarctation of the aorta, ascending aneurysm	XO mice (with a single X chromosome): no phenotypic heart disease	Unknown ⁹⁷
Chemical model			
	No human phenotype	Ang-II-infusion model	Ang II, MCP-1, IL-6, TGF- β ⁹⁸
	No human phenotype	Elastase-infusion model	Unknown ⁹⁹
	No human phenotype	Periarterial calcium application	JNK1 (ref. 100)

E10.5, embryonic day 10.5; EDS, Ehlers–Danlos syndrome; KO, knockout; LOF, loss of function; mTOR, mammalian target of rapamycin; PCP, planar cell polarity. *PTPN11* is also known as *SHP2*.

developing autopod, fibrillin-2-deficient mice show BMP7 deficiency and recapitulate the syndactyly phenotype observed in BMP7-targeted mice²⁰. Mice deficient in both fibrillin-1 and fibrillin-2 show persistent truncus arteriosus, which historically has been associated with loss of TGF- β signalling (F. Ramirez, personal communication).

Downstream of TGF- β

Little is known about the precise pathogenetic sequence downstream of TGF- β that is involved in aneurysm progression. Enhancement of matrix metalloproteinase (MMP) activity is frequently invoked. Such a model is both theoretically appealing and experimentally validated. Evidence includes high levels of MMP expression and activity in many natural and experimentally induced presentations of aneurysm, and the ability of MMP inhibitors (such as doxycycline) to attenuate aneurysm progression, including in MFS mouse models²¹. Although TGF- β has

been associated with reduced expression and activity of several MMPs in many tissues and contexts, it has been shown to specifically induce MMP2 and MMP9 expression — the MMPs that are most closely associated with aneurysm conditions such as MFS^{22,23}.

Most studies of TGF- β -related disease states have focused on ‘canonical’ (SMAD-dependent) signalling cascades, with a more historic than empirical basis for such an emphasis. More recently, it has been shown that ligand-activated TGF- β receptors can stimulate signalling through non-canonical pathways, including the phosphatidylinositol-3-OH kinase (PI(3)K)/AKT cascade, the Rho-associated protein kinase (ROCK) cascade and the mitogen-activated protein kinase (MAPK) cascade²⁴ (Fig. 2). Although mouse and human models of MFS demonstrate upregulation of canonical signalling⁷, its importance is not clear. There is emerging evidence that MAPKs may have a role in aneurysm. For example, activation of p38 has been observed in the aorta

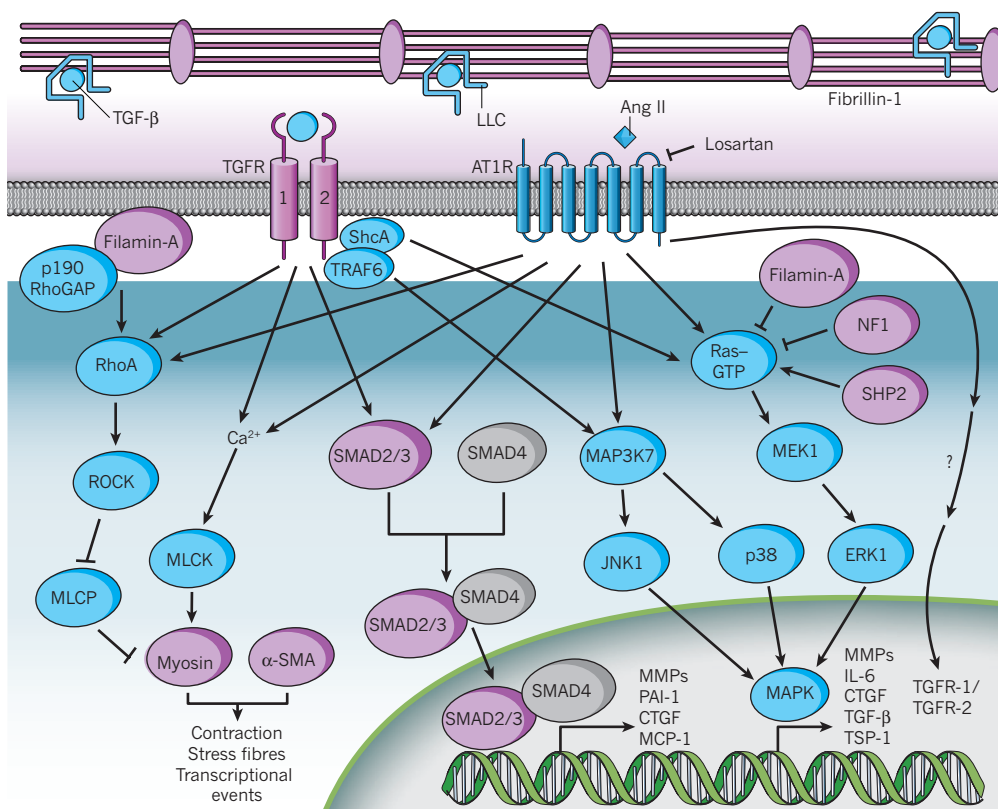


Figure 2 | The TGF- β and Ang II signalling pathways. Signalling cascades of TGF- β and Ang II have been implicated in the pathogenesis of aneurysm. Fibrillin-1, the major component of extracellular microfibrils, binds and sequesters the large latent complex (LLC) of TGF- β . After TGF- β activation (release), ligand binds to the TGF- β receptor (TGF- β R) and activates both canonical (grey) and non-canonical (blue) signalling cascades. The extensive crosstalk between the TGF- β and Ang II type 1 receptor (AT1R) signalling pathways is indicated. Key terminal events in the pathogenesis of aneurysm may include MMP-mediated proteolysis, CTGF-mediated epithelial-to-mesenchymal transition and tissue remodelling, or IL-6- and MCP-1-mediated inflammation. Proteins indicated in purple have been directly implicated in human hereditary aneurysmal disease (see Table 1). MAP3K7, mitogen-activated protein kinase kinase 7 (also known as TAK1); MEK1, MAP kinase kinase 1; MLCK, myosin light chain kinase; MLCP, myosin light chain phosphatase; p190 RhoGAP, Rho GTPase-activating protein 5; SHP2, protein tyrosine phosphatase 2C; α -SMA, α -smooth muscle actin.

of young mice homozygous for a hypomorphic fibrillin-1 (*Fbn1*) allele²⁵. We have also observed TGF- β - and angiotensin II type 1 receptor (AT1R)-dependent activation of the extracellular-signal regulated kinases (ERK1 and ERK2) in the aorta of fibrillin-1-deficient mice, and abrogation of pathological aortic root growth after treatment with a specific ERK inhibitor^{26,27}. ERK1 upregulation has also been observed in fibulin-4 deficiency in mice and humans^{4,13}, perhaps providing a link between the loss of function of fibulin-4 and fibrillin-1.

Infusion of angiotensin II (Ang II) or the application of CaCl_2 promotes AAA in mice. Antagonism of c-Jun N-terminal protein kinase 1 (JNK1) signalling has been shown to attenuate disease progression in an AAA mouse model²⁸. This occurred in association with reduced MMP2 and MMP9 activity and enhanced LOX expression. ERK activity is known to be instrumental in diverse aspects of vascular pathology, including VSMC proliferation and migration, and has been linked to TGF- β -mediated MMP upregulation and epithelial-to-mesenchymal transition. Finally, low penetrance aneurysm has been observed in human conditions known to modify Ras signalling, a major upstream activator of ERK. These include gain-of-function mutations in the *PTPN11* gene, which encodes a protein tyrosine phosphatase, leading to Noonan syndrome, and loss-of-function mutations in the *NF1* gene, which encodes a Ras GTPase-activating protein, leading to neurofibromatosis type 1 (refs 29, 30) (Table 1 and Fig. 2).

Angiotensin II and aneurysm

Aortic aneurysm and dissection can be modelled through the infusion of Ang II in mice deficient for apolipoprotein E (*Apoe*^{-/-}) (ref. 4 and references therein), or with higher doses in aged wild-type mice³¹. Aneurysm formation occurs with high penetrance in the suprarenal abdominal aorta; the ascending aorta can also be involved with lower frequency and severity. Aneurysm has been shown to be independent of hypercholesterolaemia and hypertension in this model, but requires intact AT1R signalling, innate immunity and MMP activity⁴.

The increased expression of monocyte chemoattractant protein-1 (MCP-1), its receptor (CCR2) and interleukin-6 (IL-6) have been demonstrated in models of Ang-II-induced aneurysm, and CCR2

signalling has been shown to contribute to IL-6 expression. Ang II infusion in mice was shown to associate with accumulation of CCR2⁺ macrophages in the vascular adventitia — the outermost vessel tissue layer — specifically at the sites of aneurysm formation and most prominently at the sites of dissection³¹. Mice lacking CCR2 protein showed reduced macrophage accumulation, decreased IL-6 and MCP-1 expression, and protection from dissection in response to Ang II infusion. *In vitro* modelling has demonstrated that monocytes co-cultured with adventitial fibroblasts upregulate IL-6 and MCP-1, and show enhanced differentiation into macrophages. Although the activity of a fibroblast-derived paracrine factor was suggested, there has been no speculation about its identity. The potent anti-inflammatory cytokine TGF- β is a promising candidate. TGF- β is known to induce the expression of both IL-6 and MCP-1 in many cell types, including fibroblasts and VSMCs, and can positively regulate monocyte recruitment and macrophage differentiation⁴. Increased MCP-1 expression was also proposed as a determinant of disease in response to JNK1 signalling, with JNK1 suppressing the expression of LOX that normally negatively regulates MCP-1 (ref. 32). It is notable that although the loss of CCR2 or IL-6 expression prevented early dissection in response to acute Ang II infusion in mice, it did not preclude dissection after chronic infusion. Increased adventitial IL-6 expression was observed in human aneurysms, but only at sites of dissection. In this light, it seems that the described IL-6 and MCP-1 amplification loop contributes to, but is not required for, Ang-II-induced aneurysm and dissection in mice, and more work needs to be done to determine its contribution to disease pathogenesis in humans.

Attempts to integrate TGF- β signalling into the pathogenesis of Ang-II-induced aneurysm models are frustrated by limited and contradictory empirical knowledge. As previously mentioned, Ang II signalling through AT1R has the capacity to enhance TGF- β signalling by inducing the expression of ligands, receptors and activators. It has also been reported that Ang II can activate the intracellular SMAD signalling cascade in VSMCs, in a TGF- β -independent manner³³ (Fig. 2). Ang II can also regulate MAPK signalling cascades independently of TGF- β , with the suggestion that signalling through its different receptor subtypes (AT1R and AT2R) can have varying and

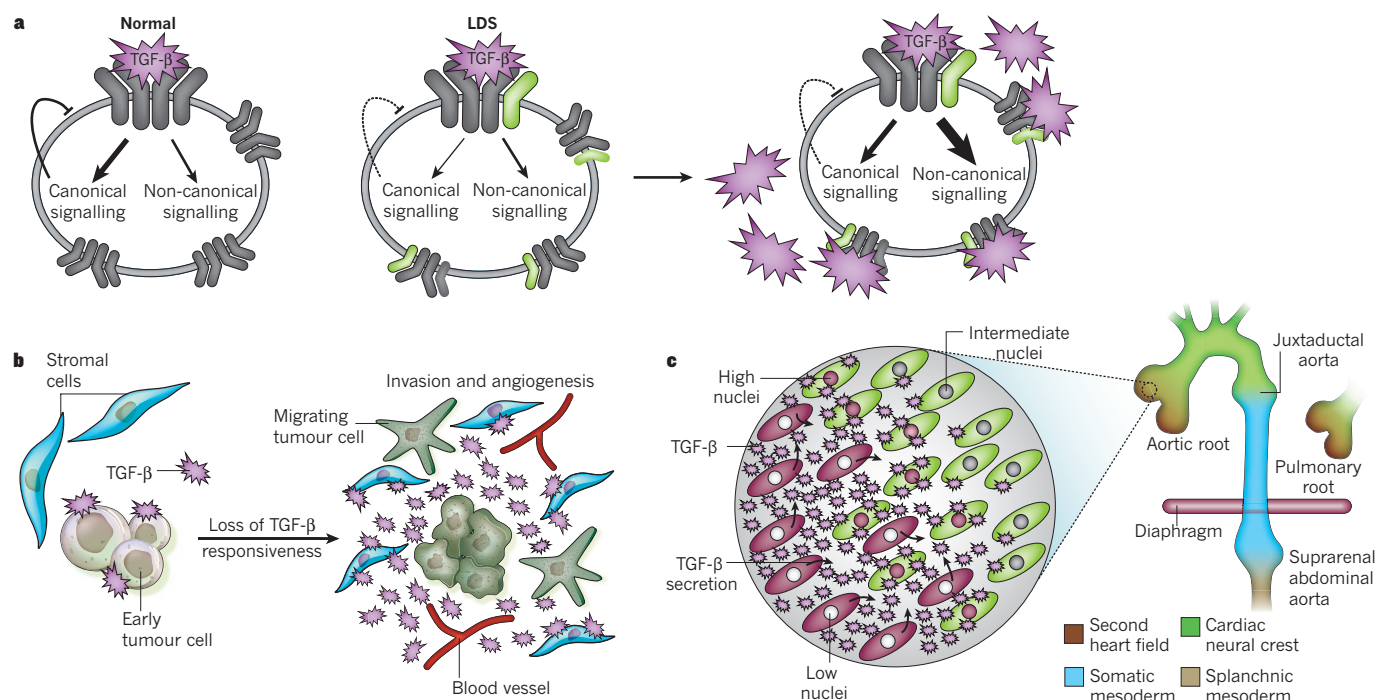


Figure 3 | TGF- β signalling in heritable aneurysm syndromes. The proposed mechanisms for the amplification of TGF- β signalling in conditions such as LDS are shown. **a**, Potential cell-autonomous mechanism of upregulation of TGF- β signalling in LDS. TGF- β ligand usually stimulates canonical and non-canonical pathways (left), with feedback inhibition provided by canonical signalling. In LDS (centre), TGF- β receptor kinase-domain mutations (depicted in green) may cause a selective decrease in canonical signalling and thus in feedback inhibition. Cell-autonomous compensation (that is, increased ligand expression and activation) to maintain canonical signalling (right) would result in excessive activation of non-canonical signalling cascades. **b**, The TGF- β cancer paradox. During tumorigenesis progression, tumour cells often lose TGF- β responsiveness as a method of escaping TGF- β -mediated cell-cycle arrest. A lack of feedback

inhibition results in upregulation of TGF- β expression by tumour cells and excessive activation of neighbouring signalling-competent stromal cells, which promotes angiogenesis and tumour invasion. **c**, Potential non-cell-autonomous mechanism of upregulation of TGF- β signalling in vascular disease. Sites of developmental field boundaries correspond anatomically to sites of predisposition for aneurysm in TGF- β vasculopathy syndromes (the aortic and pulmonary roots, the juxtaductal aorta and the suprarenal abdominal aorta). Inset shows cellular events thought to occur at the transition between second heart field (brown)- and cardiac neural crest (green)-derived VSMCs. A relative perturbation of TGF- β signalling would have a disproportionate effect on the more vulnerable lineage (second heart field), resulting in increased ligand expression and excessive TGF- β signalling by adjacent cells of a different lineage (cardiac neural crest) with relative preservation of signalling potential.

even opposing effects. Conversely, a recent study has shown that the resistance to Ang-II-induced aneurysm in normocholesterolaemic C57BL/6 mice is disrupted by systemic treatment with a neutralizing anti-TGF- β antibody³⁴. Again, these lesions were inflammatory in nature, and the incidence of aneurysm and dissection was greatly attenuated after monocyte depletion. These observations indicate that TGF- β may be protective specifically in the setting of acute and intense inflammation. Another study has shown that neutralizing anti-TGF- β antibody provided significant protection from Ang-II-induced inflammatory aneurysms after targeted silencing of CXCL10, a known chemoattractant for monocytes and macrophages³⁵. Taken together, these data suggest that TGF- β has biphasic and discordant roles in the pathogenesis of mouse Ang-II-induced aneurysm, and that TGF- β antagonism can be protective in a context-dependent manner.

Smooth muscle cytoskeletal elements and aneurysm

Studies with MFS mouse models have shown that phenotypic changes in aortic VSMCs precede elastolysis and gross medial remodelling²². These changes include adoption of a general 'synthetic' (as opposed to 'contractile') character, and morphological changes consistent with cytoskeletal rearrangement. More recent work identifying genes associated with isolated familial TAA has directly implicated perturbation of the contractile apparatus in the pathogenesis of aneurysm. Heterozygous mutations in *MYH11*, which encodes smooth muscle myosin heavy chain 11 (MYH11), cause non-syndromic ascending aortic aneurysm in association with patent ductus arteriosus and rare incidence of bicuspid aortic valve^{36,37}. By contrast, heterozygous

mutations in *ACTA2*, which encodes α -smooth muscle actin (α -SMA), cause a vascular disorder involving high, but incomplete, penetrance of ascending aortic aneurysm and dissection, with a lower incidence of descending aortic aneurysm and dissection, patent ductus arteriosus and bicuspid aortic valve³⁸. Many patients show a purplish discoloration of the skin in a network pattern, caused by altered tone in deep dermal capillaries (livedo reticularis) and pigmented cysts of the iris (iris flocculi). VSMCs isolated from patients with *MYH11* mutations showed high proliferative rates, and upregulation of insulin-like growth factor 1 (IGF-1) signalling³⁷ and components of the Ang II signalling cascade. Although it has been reported that isolated patient cells do not show evidence of increased TGF- β signalling, the data and experimental conditions were not reported. This is particularly important, because the high TGF- β signalling seen in the aorta of humans and/or mice with MFS or LDS is not recapitulated in cultured VSMCs, suggesting the necessity of tissue-specified contexts. More recently, a study found that aortic tissue from a patient with *MYH11* mutations showed high VSMC expression of the calcium-binding protein S100A12, an event previously linked to high TGF- β expression and signalling in mice³⁹.

In addition to *MYH11* and α -SMA, a third cytoskeletal (but non-contractile) protein has been implicated in aortic aneurysm. The large actin-binding protein filamin-A shows altered expression or function in the neurological condition periventricular nodular heterotopia⁴⁰. Encoded by the *FLNA* gene, on the X chromosome, filamin-A has a diverse repertoire of binding partners, making the delineation of a specific causal pathway difficult. Moreover, only a small subset of filamin-A-deficient women have been reported to show a predisposition

for aneurysm, generally in association with other systemic connective tissue findings. Mutations in *FLNA* have independently been implicated in myxomatous valve dystrophy, a phenotype commonly seen in syndromic aneurysm conditions such as MFS and LDS. Filamin-A can function as a positive regulator of TGF- β signalling through modulation of RhoA and SMAD protein trafficking, but also contributes to negative regulation of ERKs^{41,42} (Fig. 2).

Many aspects of TGF- β signalling have links to the cytoskeleton, prominently including trafficking and activity of TGF- β receptors and signalling effectors (reviewed in ref. 24). Filamin-A can bind to receptor-activated SMAD proteins, including SMAD2, and filamin-A-deficient melanoma cells show impaired TGF- β signalling compared with cells transfected with a filamin-A-encoding vector⁴². The force generated by cellular contraction against a resistance imposed by the neighbouring matrix has been shown to contribute positively to TGF- β activation in an integrin-dependent manner⁴³. Acute disruption of the actin cytoskeleton in human mesangial cells using cytochalasin D, for example, has been shown to reduce SMAD2 phosphorylation and TGF- β -induced collagen α -1(I) (but not α -1(IV)) expression⁴⁴. Other studies have shown that cytoskeletal disruption can induce the expression of TGF- β -driven gene products, such as plasminogen activator inhibitor-1 (PAI-1) and connective tissue growth factor, in VSMCs independently of TGF- β ligand or receptor, and that these effects are at least partly mediated by activation of the ROCK and/or MAPK signalling cascades⁴⁵. It remains impossible to determine the chronic consequences of such manipulations in the context of a healthy or diseased tissue.

Prospects for aneurysm treatment

Standard medical therapy for aortic aneurysm has revolved around blood pressure control to limit aortic wall stress. The implication of TGF- β signalling in the pathogenesis of aortic aneurysm suggested an opportunity for more specific therapy. Blockade of Ang II signalling through AT1R had previously been shown to limit TGF- β signalling and fibrosis in rodent models of chronic kidney disease. Indeed, in a mouse model of MFS, the AT1R blocker losartan prevented progressive aortic aneurysm⁷. Potential mechanisms for aneurysm treatment include the prevention of AT1R-induced expression of TGF- β ligands, receptors and activators such as thrombospondin-1 (TSP-1) or MMPs. Prenatal initiation of losartan treatment in MFS mouse models resulted in full normalization of aortic root size, aortic root growth rate and aortic wall architecture. Importantly, postnatal initiation of therapy in the context of established aneurysmal dilatation and medial degeneration also achieved full suppression of aortic root growth and productive remodelling of the aortic wall, with decreased elastin fragmentation and matrix deposition. Several observations pointed to TGF- β antagonism as the relevant mechanism. First, the protection achieved by losartan correlated with reduced nuclear accumulation of pSMAD2 and reduced expression of TGF- β -driven gene products such as PAI-1 and CTGF. Second, other agents with comparable blood-pressure-lowering effects that did not alter TGF- β signalling were associated with a small decline in aortic root growth rates compared with losartan, and had no effect on aortic wall architecture. Third, losartan limited the growth of the aortic root, which showed pathological dilatation and increased TGF- β signalling, but had no effect on the growth of other aortic segments, which showed neither. Losartan also limited aortic root growth in a subset of children with severe and rapidly progressive MFS⁴⁶. Several large and randomized clinical trials of losartan in MFS are under way⁴⁷.

Embracing paradox

Despite recent advances in understanding aneurysm pathogenesis, many paradoxes remain to be reconciled. For example, it is not clear why defects in structural or regulatory matrix elements, signalling molecules or contractile proteins culminate in focal aneurysms rather than a diffusely fragile and dilated arterial tree. If haemodynamic stress is the answer, it remains unclear why lesions occur on both the high- and low-pressure (that is, the root of the pulmonary artery) sides of the

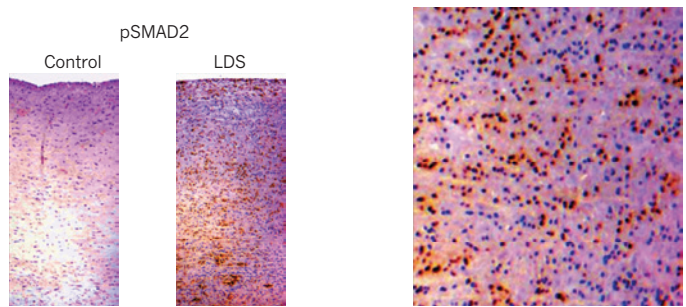


Figure 4 | TGF- β signalling in LDS aorta. Immunohistochemical staining for nuclear pSMAD2 (a marker of TGF- β signalling) in the aortic root of a control individual and a patient with LDS. An enlarged view of the LDS aorta is also shown (far right). Although TGF- β signalling is markedly increased in LDS, two distinct cell populations are observed in the aortic media, with absent (blue nuclei) or strongly positive (brown nuclei) activity. Original magnifications, $\times 10$ (left two images; from a montage of images spanning the entire thickness of the aortic wall) and $\times 60$ (far right image).

circulation. Another paradox is why conditions such as MFS and LDS show a high predisposition for focal aneurysm and primary dissection in the periductal region of the proximal descending thoracic aorta. In such conditions, it is not known why therapies aimed at blunting TGF- β signalling at the level of ligand bioavailability (such as a neutralizing anti-TGF- β antibody or losartan) or events far downstream of TGF- β (such as MMP inhibition with doxycycline) make things better, whereas events or manipulations that target the intracellular signalling cascade (such as *TGFBR1* or *TGFBR2* mutations in LDS, or the introduction of SMAD4 haploinsufficiency in fibrillin-1-deficient mice^{26,27}) routinely make things worse. The paradoxical increase in TGF- β signalling observed in patients with LDS could be explained if receptor-mediated activation of the canonical and non-canonical cascades relies on separable activities of the receptor complex. It has been suggested that distinct kinase activities underlie the phosphorylation of SMAD proteins and the ShcA adaptor protein — the proximal events in the ERK cascade⁴⁸. Furthermore, TGF- β -mediated activation of other MAPKs has been linked to receptor-mediated ubiquitylation of the TRAF6 ubiquitin ligase⁴⁹ (Fig. 2). If abnormal receptor complexes in LDS have relative preservation of non-canonical signalling but feedback regulation is disproportionately governed by canonical signalling, then compensatory mechanisms, such as increased ligand expression and/or activation, would drive excessive non-canonical TGF- β signalling in a cell-autonomous manner (Fig. 3a).

Cancer biology as a guide to aneurysm

In many respects, there are intriguing parallels to be drawn with the TGF- β cancer paradox. TGF- β can function as a tumour suppressor, with prominent roles in the maintenance of cellular differentiation and induction of cell-cycle arrest and apoptosis. During early tumorigenesis, many tumour types lose responsiveness to TGF- β through biallelic loss-of-function mutations in genes that encode the TGF- β receptors or intracellular mediators of signalling (ref. 4 and references therein). Attenuation or loss of tumour responsiveness to TGF- β leads to increased signalling by the neighbouring signalling-competent stroma, owing to increased TGF- β ligand expression (Fig. 3b). Consequences include impaired tumour surveillance owing to inhibition of adaptive immunity, acceleration of tumour growth because of enhanced angiogenesis, tumour invasion and metastasis, enhancement of innate immunity (mediated, at least in part, through MCP-1) and/or stimulation of epithelial-to-mesenchymal transition. TGF- β signalling can be further amplified in the tumour microenvironment through enhanced ligand expression by recruited inflammatory cells or enhanced TGF- β activation as a consequence of increased expression of activators such as MMPs. This sequence of events is not simply a function of the cell types per se, but more crucially the interface between cell types with a

mismatch in their intrinsic capacity to support TGF- β signalling⁵⁰.

Could the cancer paradox inform the pathogenesis of aneurysm? The first potential link comes from a consideration of cellular ontology in the vascular system. VSMCs at the base of the aorta and pulmonary artery are derived from specialized cardiogenic mesoderm termed the second heart field (Fig. 3c), a finding with possible implications for MFS⁵¹. The ascending aorta is a chimaera between cells derived from the second heart field and the ectodermal cardiac neural crest (CNC), whereas the CNC is the sole origin of VSMCs in the more distal ascending aorta and the transverse arch. There is an abrupt transition to somatic mesoderm-derived cells in the proximal descending thoracic (juxtaductal) aorta and a contribution of splanchnic mesoderm to the descending aorta beginning just below the diaphragm⁵². Thus, although there is not a common origin for VSMCs at frequent sites of predisposition for aneurysm, an interaction between cells of divergent origin exists at each location. Lineage-specific differences in intrinsic TGF- β signalling capacity were addressed in a study comparing the performance of ectoderm-derived cells from the aortic arch and mesoderm-derived cells from the abdominal aorta of chick embryos⁵³. After stimulation with TGF- β 1, ectoderm-derived VSMCs showed increased DNA synthesis and robust transcriptional activation of a TGF- β reporter allele, whereas mesoderm-derived VSMCs showed little reporter activation and growth inhibition. TGF- β responsiveness correlated with differences in the glycosylation status of TGF- β 2. The authors concluded that different SMC populations within a common vessel wall respond in lineage-dependent ways to growth factors that govern developmental events and that might participate in vascular disease in later life⁵³.

It seems plausible that cells of one lineage, such as mesodermal cells in the aortic root and descending thoracic aorta, would be more sensitive to a perturbation of TGF- β signalling. For example, this could be caused by a failure to concentrate cytokine in the setting of fibrillin-1 deficiency or by altered signalling capacity owing to heterozygous loss-of-function TGF- β receptor mutations. Loss of TGF- β feedback would initiate compensatory events such as increased TGF- β ligand expression, which could, in turn, stimulate neighbouring cells that can tolerate the primary insult better owing to improved reserves in their inherent signalling capacity (such as CNC-derived VSMCs). Indeed, inspection of the aortic wall from aneurysm tissue obtained at surgery shows an apparent binary status of medial VSMCs for TGF- β signalling (as determined by staining for nuclear pSMAD2), with neighbouring cells showing either strong or absent activity (Fig. 4). Such a model not only accommodates but also mandates a mechanism for impaired TGF- β signalling, and would explain how compensatory events could lead to functional overshoot. It would also explain why aneurysms occur at the margins of the CNC developmental field, but not typically in its middle (Figs 1 and 3c). In such a model, manipulations that limit TGF- β ligand bioavailability or block terminal pathogenetic events would be effective at preventing aneurysm, whereas those that accentuate the signalling imbalance would prove detrimental. Despite challenges, the effort to refine mechanistic understanding is justified by the high probability that these insights will yield treatment strategies for aneurysms and perhaps other clinical states associated with impaired vessel wall homeostasis. ■

1. Szabo, Z. *et al.* Aortic aneurysmal disease and cutis laxa caused by defects in the elastin gene. *J. Med. Genet.* **43**, 255–258 (2006).
2. Loeys, B. *et al.* Homozygosity for a missense mutation in fibulin-5 (*FBLN5*) results in a severe form of cutis laxa. *Hum. Mol. Genet.* **11**, 2113–2118 (2002).
3. Dasouki, M. *et al.* Compound heterozygous mutations in fibulin-4 causing neonatal lethal pulmonary artery occlusion, aortic aneurysm, arachnodactyly, and mild cutis laxa. *Am. J. Med. Genet. A* **143A**, 2635–2641 (2007).
4. Dietz, H. C. TGF- β in the pathogenesis and prevention of disease: a matter of aneurysmic proportions. *J. Clin. Invest.* **120**, 403–407 (2010).
5. Bergen, A. A. *et al.* Mutations in *ABCC6* cause pseudoxanthoma elasticum. *Nature Genet.* **25**, 228–231 (2000).
6. Neptune, E. R. *et al.* Dysregulation of TGF- β activation contributes to pathogenesis in Marfan syndrome. *Nature Genet.* **33**, 407–411 (2003).
This study first implicated upregulation of TGF- β activity in the multisystem manifestations of Marfan syndrome.
7. Habashi, J. P. *et al.* Losartan, an AT1 antagonist, prevents aortic aneurysm in a mouse model of Marfan syndrome. *Science* **312**, 117–121 (2006).
8. Loeys, B. L. *et al.* A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in *TGFBR1* or *TGFBR2*. *Nature Genet.* **37**, 275–281 (2005).
9. Mizuguchi, T. *et al.* Heterozygous *TGFBR2* mutations in Marfan syndrome. *Nature Genet.* **36**, 855–860 (2004).
10. Loeys, B. L. *et al.* Aneurysm syndromes caused by mutations in the TGF- β receptor. *N. Engl. J. Med.* **355**, 788–798 (2006).
11. van de Laar, I. M. B. H. *et al.* Mutations in *SMAD3* cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. *Nature Genet.* **43**, 121–126 (2010).
12. Coucke, P. J. *et al.* Mutations in the facilitative glucose transporter GLUT10 alter angiogenesis and cause arterial tortuosity syndrome. *Nature Genet.* **38**, 452–457 (2006).
13. Huang, J. *et al.* Fibulin-4 deficiency results in ascending aortic aneurysms: a potential link between abnormal smooth muscle cell phenotype and aneurysm progression. *Circ. Res.* **106**, 583–592 (2010).
14. Horiguchi, M. *et al.* Fibulin-4 conducts proper elastogenesis via interaction with cross-linking enzyme lysyl oxidase. *Proc. Natl Acad. Sci. USA* **106**, 19029–19034 (2009).
This study shows fibulin-4-dependent recruitment of lysyl oxidase to elastin.
15. Maki, J. M. *et al.* Inactivation of the lysyl oxidase gene *Lox* leads to aortic aneurysms, cardiovascular dysfunction, and perinatal death in mice. *Circulation* **106**, 2503–2509 (2002).
16. Atsawasuwan, P. *et al.* Lysyl oxidase binds transforming growth factor- β and regulates its signaling via amine oxidase activity. *J. Biol. Chem.* **283**, 34229–34240 (2008).
This biochemical analysis demonstrates direct inactivation of TGF- β ligand by the enzymatic activity of lysyl oxidase.
17. Zacchigna, L. *et al.* Emilin1 links TGF- β maturation to blood pressure homeostasis. *Cell* **124**, 929–942 (2006).
18. Choudhary, B. *et al.* Absence of TGF- β signaling in embryonic vascular smooth muscle leads to reduced lysyl oxidase expression, impaired elastogenesis, and aneurysm. *Genesis* **47**, 115–121 (2009).
19. Langlois, D. *et al.* Conditional inactivation of TGF- β type II receptor in smooth muscle cells and epicardium causes lethal aortic and cardiac defects. *Transgenic Res.* **19**, 1069–1082 (2010).
20. Arteaga-Solis, E. *et al.* Regulation of limb patterning by extracellular microfibrils. *J. Cell Biol.* **154**, 275–281 (2001).
This study demonstrates that fibrillins can act as positive modulators of TGF- β superfamily function.
21. Chung, A. W., Yang, H. H., Radomski, M. W. & van Breemen, C. Long-term doxycycline is more effective than atenolol to prevent thoracic aortic aneurysm in Marfan syndrome through the inhibition of matrix metalloproteinase-2 and -9. *Circ. Res.* **102**, e73–e85 (2008).
22. Bunton, T. E. *et al.* Phenotypic alteration of vascular smooth muscle cells precedes elastolysis in a mouse model of Marfan syndrome. *Circ. Res.* **88**, 37–43 (2001).
23. Sakalihasan, N., Delvenne, P., Nusgens, B. V., Limet, R. & Lapière, C. M. Activated forms of MMP2 and MMP9 in abdominal aortic aneurysms. *J. Vasc. Surg.* **24**, 127–133 (1996).
24. Moustakas, A. & Heldin, C.-H. The regulation of TGF β signal transduction. *Development* **136**, 3699–3714 (2009).
25. Carta, L. *et al.* p38 MAPK is an early determinant of promiscuous Smad2/3 signaling in the aortas of fibrillin-1 (*Fbn1*)-null mice. *J. Biol. Chem.* **284**, 5630–5636 (2009).
26. Holm, T. *et al.* Noncanonical TGF- β signaling contributes to aortic aneurysm progression in Marfan syndrome mice. *Science* **332**, 358–361 (2011).
27. Habashi, J. P. *et al.* Angiotensin II type 2 receptor signaling attenuates aortic aneurysm in mice through ERK antagonism. *Science* **332**, 361–365 (2011).
28. Yoshimura, K. *et al.* Regression of abdominal aortic aneurysm by inhibition of c-Jun N-terminal kinase. *Nature Med.* **11**, 1330–1338 (2005).
29. Purnell, R., Williams, I., Von Oppell, U. & Wood, A. Giant aneurysms of the sinuses of Valsalva and aortic regurgitation in a patient with Noonan's syndrome. *Eur. J. Cardiothorac. Surg.* **28**, 346–348 (2005).
30. Friedman, J. M. *et al.* Cardiovascular disease in neurofibromatosis 1: report of the NF1 Cardiovascular Task Force. *Genet. Med.* **4**, 105–111 (2002).
31. Tieu, B. C. *et al.* An adventitial IL-6/MCP1 amplification loop accelerates macrophage-mediated vascular inflammation leading to aortic dissection in mice. *J. Clin. Invest.* **119**, 3637–3651 (2009).
This paper describes a major role for adventitial macrophage in aortic dissection in mice.
32. Onoda, M. *et al.* Lysyl oxidase resolves inflammation by reducing monocyte chemoattractant protein-1 in abdominal aortic aneurysm. *Atherosclerosis* **208**, 366–369 (2010).
33. Rodríguez-Vita, J. *et al.* Angiotensin II activates the Smad pathway in vascular smooth muscle cells by a transforming growth factor- β -independent mechanism. *Circulation* **111**, 2509–2517 (2005).
34. Wang, Y. *et al.* TGF- β activity protects against inflammatory aortic aneurysm progression and complications in angiotensin II-infused mice. *J. Clin. Invest.* **120**, 422–432 (2010).
35. King, V. L. *et al.* Interferon- γ and the interferon-inducible chemokine CXCL10 protect against aneurysm formation and rupture. *Circulation* **119**, 426–435 (2009).
36. Zhu, L. *et al.* Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. *Nature Genet.* **38**, 343–349 (2006).
This was the first study to implicate the smooth muscle contraction apparatus in thoracic aortic aneurysm.
37. Pannu, H. *et al.* *MYH11* mutations result in a distinct vascular pathology driven by insulin-like growth factor 1 and angiotensin II. *Hum. Mol. Genet.* **16**, 2453–2462 (2007).

38. Guo, D.-C. *et al.* Mutations in smooth muscle α -actin (*ACTA2*) lead to thoracic aortic aneurysms and dissections. *Nature Genet.* **39**, 1488–1493 (2007).
39. Hofmann Bowman, M. *et al.* S100A12 mediates aortic wall remodeling and aortic aneurysm. *Circ. Res.* **106**, 145–154 (2010).
40. Sheen, V. L. *et al.* Filamin A mutations cause periventricular heterotopia with Ehlers–Danlos syndrome. *Neurology* **64**, 254–262 (2005).
41. Zhu, T.-N. *et al.* Filamin A-mediated down-regulation of the exchange factor Ras-GRF1 correlates with decreased matrix metalloproteinase-9 expression in human melanoma cells. *J. Biol. Chem.* **282**, 14816–14826 (2007).
42. Sasaki, A., Masuda, Y., Ohta, Y., Ikeda, K. & Watanabe, K. Filamin associates with Smads and regulates transforming growth factor- β signaling. *J. Biol. Chem.* **276**, 17871–17877 (2001).
43. Wipff, P.-J., Rifkin, D. B., Meister, J.-J. & Hinz, B. Myofibroblast contraction activates latent TGF- β 1 from the extracellular matrix. *J. Cell Biol.* **179**, 1311–1323 (2007).
44. Hubchak, S. C., Runyan, C. E., Kreisberg, J. I. & Schnaper, H. W. Cytoskeletal rearrangement and signal transduction in TGF- β 1-stimulated mesangial cell collagen accumulation. *J. Am. Soc. Nephrol.* **14**, 1969–1980 (2003).
45. Samarakoon, R., Higgins, C. E., Higgins, S. P. & Higgins, P. J. Differential requirement for MEK/ERK and SMAD signaling in PAI-1 and CTGF expression in response to microtubule disruption. *Cell. Signal.* **21**, 986–995 (2009).
46. Brooke, B. S. *et al.* Angiotensin II blockade and aortic-root dilation in Marfan's syndrome. *N. Engl. J. Med.* **358**, 2787–2795 (2008).
47. Lacro, R. V. *et al.* Rationale and design of a randomized clinical trial of β -blocker therapy (atenolol) versus angiotensin II receptor blocker therapy (losartan) in individuals with Marfan syndrome. *Am. Heart J.* **154**, 624–631 (2007).
48. Lee, M. K. *et al.* TGF- β activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *EMBO J.* **26**, 3957–3967 (2007).
49. Yamashita, M. *et al.* TRAF6 mediates Smad-independent activation of JNK and p38 by TGF- β . *Mol. Cell* **31**, 918–924 (2008).
50. Bachman, K. E. & Park, B. H. Dual nature of TGF- β signaling: tumor suppressor vs. tumor promoter. *Curr. Opin. Oncol.* **17**, 49–54 (2005).
51. Waldo, K. L. *et al.* Secondary heart field contributes myocardium and smooth muscle to the arterial pole of the developing heart. *Dev. Biol.* **281**, 78–90 (2005).
52. Majesky, M. W. Developmental basis of vascular smooth muscle diversity. *Arterioscler. Thromb. Vasc. Biol.* **27**, 1248–1258 (2007).
53. Topouzis, S. & Majesky, M. W. Smooth muscle lineage diversity in the chick embryo: two types of aortic smooth muscle cell differ in growth and receptor-mediated transcriptional responses to transforming growth factor- β . *Dev. Biol.* **178**, 430–445 (1996).
- This study notes divergent TGF- β responsiveness in different aortic regions.**
54. Carta, L. *et al.* Fibrillins 1 and 2 perform partially overlapping functions during aortic development. *J. Biol. Chem.* **281**, 8016–8023 (2006).
55. Dietz, H. C. *et al.* Marfan syndrome caused by a recurrent *de novo* missense mutation in the fibrillin gene. *Nature* **352**, 337–339 (1991).
56. Li, D. Y. *et al.* Elastin is an essential determinant of arterial morphogenesis. *Nature* **393**, 276–280 (1998).
57. Rahkonen, O. *et al.* Mice with a deletion in the first intron of the *Col1a1* gene develop age-dependent aortic dissection and rupture. *Circ. Res.* **94**, 83–90 (2004).
58. Malfait, F. *et al.* Three arginine to cysteine substitutions in the pro- α 1(I)-collagen chain cause Ehlers–Danlos syndrome with a propensity to arterial rupture in early adulthood. *Hum. Mutat.* **28**, 387–395 (2007).
59. Schwarze, U. *et al.* Rare autosomal recessive cardiac valvular form of Ehlers–Danlos syndrome results from mutations in the *COL1A2* gene that activate the nonsense-mediated RNA decay pathway. *Am. J. Hum. Genet.* **74**, 917–930 (2004).
60. Liu, X., Wu, H., Byrne, M., Krane, S. & Jaenisch, R. Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development. *Proc. Natl Acad. Sci. USA* **94**, 1852–1856 (1997).
61. Superti-Furga, A., Gugler, E., Gitzelmann, R. & Steinmann, B. Ehlers–Danlos syndrome type IV: a multi-exon deletion in one of the two *COL3A1* alleles affecting structure, stability, and processing of type III procollagen. *J. Biol. Chem.* **263**, 6226–6232 (1988).
- This was the first genetic study to link a single-gene disorder to aneurysmal disease.**
62. Plaisier, E. *et al.* *COL4A1* mutations and hereditary angiodysplasia, nephropathy, aneurysms, and muscle cramps. *N. Engl. J. Med.* **357**, 2687–2695 (2007).
63. Pöschl, E. *et al.* Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development* **131**, 1619–1628 (2004).
64. Kashtan, C. E. *et al.* Aortic abnormalities in males with Alport syndrome. *Nephrol. Dial. Transplant.* **25**, 3554–3560 (2010).
65. Wenstrup, R. J., Murad, S. & Pinnell, S. R. Ehlers–Danlos syndrome type VI: clinical manifestations of collagen lysyl hydroxylase deficiency. *J. Pediatr.* **115**, 405–409 (1989).
66. Takaluoma, K. *et al.* Tissue-specific changes in the hydroxylysine content and cross-links of collagens and alterations in fibril morphology in lysyl hydroxylase 1 knock-out mice. *J. Biol. Chem.* **282**, 6588–6596 (2007).
67. Salo, A. M. *et al.* A connective tissue disorder caused by mutations of the lysyl hydroxylase 3 gene. *Am. J. Hum. Genet.* **83**, 495–503 (2008).
68. Ruotsalainen, H. *et al.* Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *J. Cell Sci.* **119**, 625–635 (2006).
69. Larsson, J. *et al.* Abnormal angiogenesis but intact hematopoietic potential in TGF- β type I receptor-deficient mice. *EMBO J.* **20**, 1663–1673 (2001).
70. Oshima, M., Oshima, H. & Taketo, M. M. TGF- β receptor type II deficiency results in defects of yolk sac hematopoiesis and vasculogenesis. *Dev. Biol.* **179**, 297–302 (1996).
71. Hsi, D. H., Ryan, G. F., Hellems, S. O., Cheeran, D. C. & Sheils, L. A. Large aneurysms of the ascending aorta and major coronary arteries in a patient with hereditary hemorrhagic telangiectasia. *Mayo Clin. Proc.* **78**, 774–776 (2003).
72. Arthur, H. M. *et al.* Endoglin, an ancillary TGF β receptor, is required for extraembryonic angiogenesis and plays a key role in heart development. *Dev. Biol.* **217**, 42–53 (2000).
73. Andersen, N. D. *et al.* Thoracic endografting in a patient with hereditary hemorrhagic telangiectasia presenting with a descending thoracic aneurysm. *J. Vasc. Surg.* **51**, 468–470 (2010).
74. Oh, S. P. *et al.* Activin receptor-like kinase 1 modulates transforming growth factor- β 1 signaling in the regulation of angiogenesis. *Proc. Natl Acad. Sci. USA* **97**, 2626–2631 (2000).
75. Cheng, C.-H. *et al.* Mutations in the *SLC2A10* gene cause arterial abnormalities in mice. *Cardiovasc. Res.* **81**, 381–388 (2009).
76. Garg, V. *et al.* Mutations in *NOTCH1* cause aortic valve disease. *Nature* **437**, 270–274 (2005).
77. Conlon, R. A., Reaume, A. G. & Rossant, J. *Notch1* is required for the coordinate segmentation of somites. *Development* **121**, 1533–1545 (1995).
78. Xue, Y. *et al.* Embryonic lethality and vascular defects in mice lacking the Notch ligand Jagged1. *Hum. Mol. Genet.* **8**, 723–730 (1999).
79. Kamath, B. M. *et al.* Vascular anomalies in Alagille syndrome: a significant cause of morbidity and mortality. *Circulation* **109**, 1354–1358 (2004).
80. Distefano, G. *et al.* Polycystin-1 regulates extracellular signal-regulated kinase-dependent phosphorylation of tuberin to control cell size through mTOR and its downstream effectors S6K and 4EBP1. *Mol. Cell. Biol.* **29**, 2359–2371 (2009).
81. Hassane, S. *et al.* Pathogenic sequence for dissecting aneurysm formation in a hypomorphic polycystic kidney disease 1 mouse model. *Arterioscler. Thromb. Vasc. Biol.* **27**, 2177–2183 (2007).
82. Wu, G. *et al.* Somatic inactivation of *Pkd2* results in polycystic kidney disease. *Cell* **93**, 177–188 (1998).
83. Yu, Q. *et al.* ENU induced mutations causing congenital cardiovascular anomalies. *Development* **131**, 6211–6223 (2004).
84. Schildmeyer, L. A. *et al.* Impaired vascular contractility and blood pressure homeostasis in the smooth muscle α -actin null mouse. *FASEB J.* **14**, 2213–2220 (2000).
85. Morano, I. *et al.* Smooth-muscle contraction without smooth-muscle myosin. *Nature Cell Biol.* **2**, 371–375 (2000).
86. Feng, Y. *et al.* Filamin A (FLNA) is required for cell–cell contact in vascular development and cardiac morphogenesis. *Proc. Natl Acad. Sci. USA* **103**, 19836–19841 (2006).
87. Araki, T. *et al.* Mouse model of Noonan syndrome reveals cell type- and gene dosage-dependent effects of *Ptpn11* mutation. *Nature Med.* **10**, 849–857 (2004).
88. Iwasaki, Y. *et al.* Coronary artery dilatation in LEOPARD syndrome. A child case and literature review. *Congenit. Heart Dis.* **4**, 38–41 (2009).
89. Takahashi, H. *et al.* A hereditary model of slowly progressive polycystic kidney disease in the mouse. *J. Am. Soc. Nephrol.* **1**, 980–989 (1991).
90. Lee, T. C., Zhao, Y. D., Courtman, D. W. & Stewart, D. J. Abnormal aortic valve development in mice lacking endothelial nitric oxide synthase. *Circulation* **101**, 2345–2348 (2000).
91. Kuhlencordt, P. J. *et al.* Accelerated atherosclerosis, aortic aneurysm formation, and ischemic heart disease in apolipoprotein E/endothelial nitric oxide synthase double-knockout mice. *Circulation* **104**, 448–454 (2001).
92. Cao, J. *et al.* Thoracic aortic disease in tuberous sclerosis complex: molecular pathogenesis and potential therapies in *Tsc2*^{−/−} mice. *Hum. Mol. Genet.* **19**, 1908–1920 (2010).
93. Raben, N. *et al.* Targeted disruption of the acid α -glucosidase gene in mice causes an illness with critical features of both infantile and adult human glycogen storage disease type II. *J. Biol. Chem.* **273**, 19086–19092 (1998).
94. Rocha, P. P., Scholze, M., Bleiss, W. & Schrewe, H. Med12 is essential for early mouse development and for canonical Wnt and Wnt/PCP signaling. *Development* **137**, 2723–2731 (2010).
95. Halder, S. M. *et al.* *Klf15* deficiency is a molecular link between heart failure and aortic aneurysm formation. *Sci. Transl. Med.* **2**, 26ra26 (2010).
96. Kuo, C. T. *et al.* The LKLF transcription factor is required for normal tunica media formation and blood vessel stabilization during murine embryogenesis. *Genes Dev.* **11**, 2996–3006 (1997).
97. Lopez, L. *et al.* Turner syndrome is an independent risk factor for aortic dilation in the young. *Pediatrics* **121**, e1622–e1627 (2008).
98. Daugherty, A., Manning, M. W. & Cassis, L. A. Angiotensin II promotes atherosclerotic lesions and aneurysms in apolipoprotein E-deficient mice. *J. Clin. Invest.* **105**, 1605–1612 (2000).
- This study formed the basis for the Ang-II-infusion model of aneurysmal disease, the most widespread mouse model of aneurysm.**
99. Anidjar, S. *et al.* Elastase-induced experimental aneurysms in rats. *Circulation* **82**, 973–981 (1990).
100. Chiou, A. C., Chiu, B. & Pearce, W. H. Murine aortic aneurysm produced by periaortic application of calcium chloride. *J. Surg. Res.* **99**, 371–376 (2001).

Acknowledgements We would like to thank E. Arbustini for the use of the multidetector computer tomographic image in Fig. 1a. We acknowledge funding support from the US National Heart, Lung, and Blood Institute, the Howard Hughes Medical Institute, the US National Marfan Foundation and the William S. Smilow Center for Marfan Research.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of this article at www.nature.com/nature. The authors declare no competing financial interests. Correspondence should be addressed to H.C.D. (hdietz@jhmi.edu).

Progress and challenges in translating the biology of atherosclerosis

Peter Libby¹, Paul M Ridker^{1,2} & Göran K. Hansson³

Atherosclerosis is a chronic disease of the arterial wall, and a leading cause of death and loss of productive life years worldwide. Research into the disease has led to many compelling hypotheses about the pathophysiology of atherosclerotic lesion formation and of complications such as myocardial infarction and stroke. Yet, despite these advances, we still lack definitive evidence to show that processes such as lipoprotein oxidation, inflammation and immunity have a crucial involvement in human atherosclerosis. Experimental atherosclerosis in animals furnishes an important research tool, but extrapolation to humans requires care. Understanding how to combine experimental and clinical science will provide further insight into atherosclerosis and could lead to new clinical applications.

Powerful laboratory research in the past decade has led to many reviews that describe the biological and genetic bases of atherosclerosis^{1–3}. Despite this progress, the leap from experimental animal findings to human atherosclerosis and clinical application presents challenges. The laboratory literature and experimental community sometimes assume that the results obtained in cultured cells or animals closely correspond to humans. Although experimental work has helped to unravel some of the principles of atherosclerosis pathophysiology, gaps remain in translation to the clinic, and these breeches require bridging to achieve the full promise of scientific advances in atherosclerosis.

This Review summarizes the burgeoning biological understanding of atherosclerosis. Instead of celebrating the astounding advances already achieved, we highlight some of the challenges to the clinical application of these advances. We also offer possible ways to move forward and overcome these obstacles.

Current concepts of atherogenesis

Atherogenesis refers to the development of atheromatous plaques in the inner lining of the arteries. On the basis of animal experiments and observations in human specimens, most contemporary schemes of atherogenesis posit an initial qualitative change in the monolayer of endothelial cells that lines the inner arterial surface (Fig. 1a). Arterial endothelial cells, which normally resist attachment of the white blood cells streaming past them, express adhesion molecules that capture leukocytes on their surfaces (Fig. 1b) when subjected to irritative stimuli (such as dyslipidaemia, hypertension or pro-inflammatory mediators). Parallel changes in endothelial permeability and the composition of the extracellular matrix beneath the endothelium promote the entry and retention of cholesterol-containing low-density lipoprotein (LDL) particles in the artery wall⁴. Biochemically modified components of these particles may induce leukocyte adhesion, and intact but modified particles undergo endocytosis by monocyte-derived macrophages, leading to intracellular cholesterol accumulation. Chemoattractant mediators direct the migration of the bound leukocytes into the innermost layer of the artery, the tunica intima (Figs 1b and 2). The localized distribution of atheromatous lesions in the arterial tree, despite a systemic rise in risk factors such as increased LDL levels or blood pressure, probably reflects differing haemodynamics in different segments of the arterial tree, distinction in the regional development

of arteries⁵ and the ability of normal laminar shear stress to elicit an atheroprotective program of gene expression by the endothelium⁶. Once resident in the artery wall, monocytes — the most numerous white blood cells in plaques — differentiate into tissue macrophages. In the nascent atheroma, these mononuclear phagocytes engulf lipoprotein particles and become foam cells — a term that reflects the microscopic appearance of these lipid-laden macrophages.

In mice, a pro-inflammatory subset of monocytes induced by hyperlipidaemia may preferentially furnish the precursors of lesional foam cells, but the fates and functions of this monocyte subset and its human equivalent remain under intense exploration^{7,8}. Macrophages in the atheroma may also have a pro-inflammatory palette of functions, characteristic of M1 macrophages⁹, which produce high levels of effectors such as the cytokines interleukin-1 β (IL-1 β) and tumour-necrosis factor (TNF). Some mononuclear phagocytes in plaques have the characteristics, and probably the antigen-presenting functions, of dendritic cells. Other leukocyte classes (such as lymphocytes) and mast cells also accumulate in atheromata, but less abundantly than phagocytes. Lesional T cells, although far fewer in number than macrophages, probably have key regulatory functions in plaques.

Atheroma formation also involves the recruitment of smooth muscle cells (SMCs) from the tunica media — the middle layer of the artery wall — into the tunica intima (Fig. 1c). Unlike that of most experimental animals used to study atherosclerosis, the intima of human arteries (including the coronary arteries) contains resident SMCs. During atherogenesis, other SMCs migrate from the media into the intima, and proliferate in response to mediators such as platelet-derived growth factor. In the intima, the SMCs produce extracellular matrix molecules, including interstitial collagen and elastin, and form a fibrous cap that covers the plaque. This cap typically overlies a collection of macrophage-derived foam cells, some of which die (for example, by apoptosis) and release lipids that accumulate extracellularly. The inefficient clearance of dead cells — a process known as efferocytosis — can promote the accumulation of cellular debris and extracellular lipids, forming a lipid-rich pool called the necrotic core of the plaque¹⁰.

Plaques generally cause clinical manifestations by producing flow-limiting stenoses that lead to tissue ischaemia, or by provoking thrombi that can interrupt blood flow locally or embolize and lodge

¹Cardiovascular Division, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA. ²Division of Cardiovascular Medicine, Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ³Department of Medicine and Center for Molecular Medicine, Karolinska University Hospital, Karolinska Institutet, Stockholm SE-17176, Sweden.

in distal arteries. Paradoxically, thrombotic complications do not always occur at the sites of the most severe arterial narrowing by plaques. Instead, thrombi often arise after physical disruption of the plaque, most commonly a fracture of the fibrous cap that exposes pro-coagulant material in the plaque's core to coagulation proteins in the blood, triggering thrombosis (Fig. 1d). Plaques that rupture typically have thin, collagen-poor fibrous caps with few SMCs but abundant macrophages. The inflammatory cells may hasten plaque disruption by elaborating collagenolytic enzymes that can degrade collagen, and by generating mediators that provoke the death of SMCs, the source of arterial collagen¹¹. Plaque macrophages also produce the pro-coagulant tissue factor that renders the lipid core thrombogenic. Thus, the infiltrating inflammatory cells interact with the intrinsic arterial cells (smooth muscle and endothelium), promoting lesion formation and complications.

The risk factors for atherosclerosis act at several points on this pathogenic pathway. Hypertension is a major risk factor for atheromata, and can increase arterial wall tension, leading to disturbed repair processes and aneurysm formation. Angiotensin II, a major

pressor hormone, can alter endothelial function, inciting leukocyte adhesion. Cigarette smoking and diabetes also affect vascular biology, but through less well understood mechanisms. The role of cholesterol has been investigated in great detail, yielding success in cardiovascular prevention strategies.

Lipids and atherosclerosis

Lipids have a central role in the pathogenesis of plaques, but the mechanistic links between lipids and atherogenesis remain unclear. Observational data support a strong association between plasma lipid levels and the risk of cardiovascular disease¹². In particular, LDL levels satisfy modified Koch's postulates — criteria for judging whether a specific microbe is the cause of a disease — for causality of atherosclerosis¹³. LDL levels correlate with the risk of cardiovascular events in human populations, and augment individual susceptibility to atherosclerosis and its complications. Monogenic disorders that raise plasma levels of LDL heighten cardiovascular risk. Several interventions that lower LDL levels by independent mechanisms diminish the likelihood of atherosclerotic events.

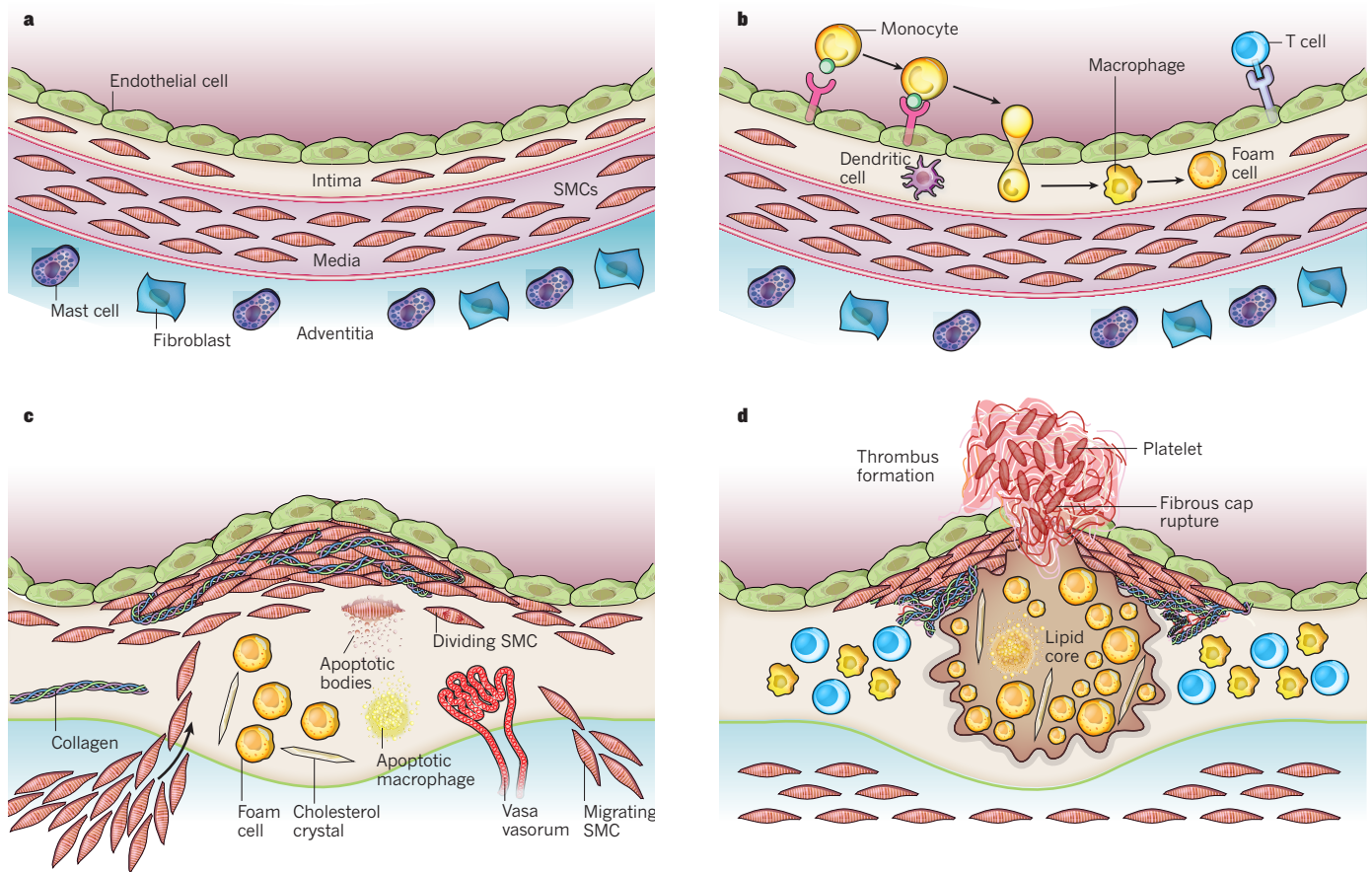


Figure 1 | Stages in the development of atherosclerotic lesions. The normal muscular artery and the cell changes that occur during disease progression to thrombosis are shown. **a**, The normal artery contains three layers. The inner layer, the tunica intima, is lined by a monolayer of endothelial cells that is in contact with blood overlying a basement membrane. In contrast to many animal species used for atherosclerosis experiments, the human intima contains resident smooth muscle cells (SMCs). The middle layer, or tunica media, contains SMCs embedded in a complex extracellular matrix. Arteries affected by obstructive atherosclerosis generally have the structure of muscular arteries. The arteries often studied in experimental atherosclerosis are elastic arteries, which have clearly demarcated laminae in the tunica media, where layers of elastin lie between strata of SMCs. The adventitia, the outer layer of arteries, contains mast cells, nerve endings and microvessels. **b**, The initial steps of atherosclerosis include adhesion of blood leukocytes to the activated endothelial monolayer, directed migration of the bound

leukocytes into the intima, maturation of monocytes (the most numerous of the leukocytes recruited) into macrophages, and their uptake of lipid, yielding foam cells. **c**, Lesion progression involves the migration of SMCs from the media to the intima, the proliferation of resident intimal SMCs and media-derived SMCs, and the heightened synthesis of extracellular matrix macromolecules such as collagen, elastin and proteoglycans. Plaque macrophages and SMCs can die in advancing lesions, some by apoptosis. Extracellular lipid derived from dead and dying cells can accumulate in the central region of a plaque, often denoted the lipid or necrotic core. Advancing plaques also contain cholesterol crystals and microvessels. **d**, Thrombosis, the ultimate complication of atherosclerosis, often complicates a physical disruption of the atherosclerotic plaque. Shown is a fracture of the plaque's fibrous cap, which has enabled blood coagulation components to come into contact with tissue factors in the plaque's interior, triggering the thrombus that extends into the vessel lumen, where it can impede blood flow.

The LDL success story lacks a final chapter

The determination of the LDL pathway and therapy with inhibitors of hydroxymethyl glutaryl coenzyme A reductase (collectively known as statins), which regulate this pathway, are conspicuous victories of cardiovascular science and medicine¹⁴. But even in patients treated with statins, a considerable residual burden of cardiovascular risk remains¹⁵. More than 20% of patients will have a recurrent event within 30 months of an acute coronary syndrome, despite receiving high-dose statin treatment¹⁶. These findings indicate that treatments to decrease LDL levels even further, beyond the targets currently mandated by various national guidelines, could provide further clinical benefit. Unfortunately, at least one-quarter of high-risk patients who receive intensive statin therapy have LDL levels above current guideline-mandated goals¹⁷. New biological targets have emerged that may yield incremental lowering of LDL levels to a greater degree than that achieved by high-dose statin therapy (Box 1).

HDL as a frustrating next frontier

Consistent evidence has shown that levels of high-density lipoprotein (HDL) correlate inversely with cardiovascular risk. Numerous approaches to increase HDL exist or are in development. Because of the heterogeneity in HDL particles, the complicated pathways of cholesterol flux mediated by HDL and the association of HDL with many proteins that may modify atherosclerosis, the steady-state levels of HDL cholesterol in blood reflect HDL function poorly. HDL particles can effect reverse cholesterol transport, and transfer cholesterol from peripheral tissues to the liver for excretion. This process involves the unloading of cholesterol from lipid-laden macrophages in atheromata by means of membrane-bound ATP-binding cassette transporters. Mature HDL interacts with one ATP-binding cassette transporter (ABCG1), and nascent HDL with another (ABCA1)^{18,19} (Fig. 2).

In addition to mediating reverse cholesterol transport, HDL can exert anti-inflammatory actions both *in vitro* and *in vivo*¹⁹. HDL particles associate with dozens of proteins, many with biological activities that have relevance to atherogenesis²⁰. The lipid content of HDL particles can be remodelled — for example, the plasma protein cholesteryl ester transfer protein (CETP) facilitates the exchange of

cholesteryl esters in HDL for triglycerides from apolipoprotein-B-containing lipoproteins^{21,22} (Fig. 2). The protein content of HDL particles can also be remodelled — for example, when plasma levels of the acute-phase reactant serum amyloid A increase during inflammatory states²³. Typical clinical assays for HDL do not reflect this high degree of heterogeneity of the particles that influence plaque biology²⁴. Thus, the mere increase in HDL levels in response to some interventions may not necessarily confer clinical benefit, owing to qualitative changes in the particles. By contrast, the lowering of LDL levels usually reduces cardiovascular event rates. Of the approaches to increase HDL under study, the potential of CETP inhibition to improve outcomes remains unclear. The CETP inhibitor torcetrapib failed in the clinic, probably owing to off-target effects^{18,25,26}, and two other CETP inhibitors, dalcetrapib and anacetrapib, have entered clinical evaluation. The safety of anacetrapib was recently affirmed by a phase III clinical trial, which provided preliminary evidence for reduced clinical events²⁷. Ultimately, the results of continuing large end-point trials should settle the CETP controversy.

Apolipoprotein A-I (Apo-AI), the major protein component of HDL, has received much attention as a possible therapeutic target for atherosclerosis^{28–30}. But difficulties have plagued the development of protein therapeutics and mimetics. Despite small biomarker studies that suggest possible efficacy of some such agents, various limitations have stalled their entry to trials that could show efficacy in cardiovascular event reduction.

Manipulation of the transcription of *APOA1* has proven elusive, with only one agent in development for this purpose. Stimulation of the nuclear receptor peroxisome proliferator-activated receptor- α (PPAR- α) moderately increases Apo-AI levels. Moreover, preclinical and biomarker studies have suggested beneficial vascular actions of PPAR- α agonism that do not depend on Apo-AI³¹. Clinical trials of one agent with PPAR- α -stimulating activity, gemfibrozil, have shown a reduction in cardiovascular events^{32,33}. Unfortunately, the combination of gemfibrozil with statins raises major safety concerns, owing to a well-defined drug–drug interaction³⁴. Another agent with relatively weak PPAR- α -agonist action, fenofibrate, has not reduced events in several large clinical trials^{35,36}.

BOX 1

New biological targets for lowering LDL levels

Here we consider biological targets that may reduce LDL levels to a greater extent than that obtained by high-dose statin therapy.

Niemann–Pick C1-like protein 1

Inhibition of the intestinal cholesterol transporter Niemann–Pick C1-like protein 1 (NPC1L1) by the agent ezetimibe can reduce LDL levels by almost 20% in individuals already being treated with statins³⁹. Although combined therapy with statins and ezetimibe can help more individuals to reach mandated LDL targets for their level of risk, no clinical trial data have so far shown that this strategy will lower cardiovascular event rates beyond the drop produced by statin monotherapy. Studies of biomarkers such as the thickness of the carotid artery intima media, flow-mediated vasodilation, or inflammation cannot supplant lacking of data on clinical events. This example emphasizes three important points: (1) the need to choose biomarkers carefully to be pursued in clinical development; (2) the ultimate requirement for clinical end-point studies to determine the efficacy of interventions; and (3) the value of starting such definitive studies early in drug-development programmes.

Proprotein convertase subtilisin/kexin type 9

Genetic studies have shown that mutations in the gene that encodes

the enzyme proprotein convertase subtilisin/kexin type 9 (PCSK9) augment LDL receptor levels on cell surfaces, boosting LDL clearance and yielding lower LDL concentrations in the blood⁴⁰. The enzymatic activity of PCSK9 — autocatalysis — does not directly degrade LDL receptors. Although enzymes generally make good drug targets, the autocatalytic activity of PCSK9 has proven difficult to inhibit by conventional medicinal chemistry approaches, and does not necessarily reflect its regulation of LDL receptor levels, spurring the development of biological agents that seek to limit PCSK9 action.

Individuals with loss-of-function variants in PCSK9, who are exposed to lower levels of LDL from childhood than those with the common genotype for this enzyme, seem protected from atherosclerotic events even when they have other cardiovascular risk factors⁴¹. This observation suggests that lowering LDL levels for longer periods than those encompassed by typical clinical trials should continue to provide benefit, and supports a pivotal, perhaps permissive, role for LDL in atherosclerosis. Such genetic data also help to clarify the importance of LDL lowering compared with other potential mechanisms of the benefits of statins (for example, statins interfere with prenylation of small G proteins, modulates lipid-raft organization and activates of Krüppel-like factor 2)^{18,92,93}.

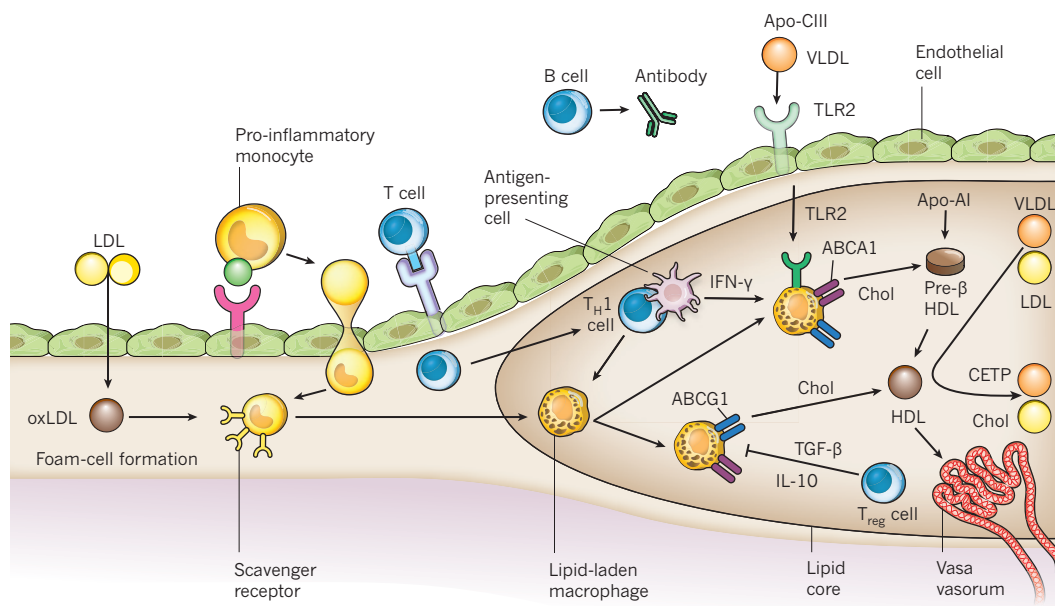


Figure 2 | The intersection of inflammation and lipid metabolism modulates atherosclerosis and provides potential targets for therapeutic manipulation. Atherogenesis begins with the recruitment of inflammatory cells to the intima. Activated endothelial cells express leukocyte adhesion molecules that capture blood monocytes, including (but not exclusively) the pro-inflammatory subset marked by high expression levels of the cell-surface protein Ly6C in mice. After inflammatory activation, monocytes recruited to the intima express scavenger receptors that permit the uptake of modified LDL particles, such as oxidized LDL (oxLDL). Cholesterol loading leads to the formation of foam cells, and ultimately leads to the mature lipid-laden macrophages of the plaque's core. These cells can produce pro-inflammatory mediators, reactive oxygen species, and tissue factor pro-coagulants that amplify local inflammation and promote thrombotic complications. Although fewer in number than the mononuclear phagocytes, T cells also enter the intima and send decisive regulatory signals. After antigen-specific activation, T helper 1 (T_H1) cells secrete the signature cytokine interferon- γ (IFN- γ), which can activate vascular wall cells and macrophages, and magnify and sustain the inflammatory response in the intima. Regulatory T (T_{reg}) cells produce interleukin-10 (IL-10) and transforming growth factor- β (TGF- β), two cytokines considered to exert anti-inflammatory actions. Although not numerically prominent in the plaque, B cells accumulate and

organize in the perivascular tissue surrounding atherosclerotic arteries. They produce circulating antibodies that may limit inflammation and mute atherogenesis. In addition to modified LDL, triglyceride-rich lipoproteins such as very-low-density lipoprotein (VLDL) — particularly those particles that bear apolipoprotein C-III (Apo-CIII) or apolipoprotein B (Apo-B) — can instigate vascular inflammation through Toll-like receptor 2 (TLR2) signalling. Macrophage foam cells can efflux cholesterol (chol) through ATP-binding cassette (ABC) transporters, which work in tandem. ABCA1 loads cholesterol-poor nascent high-density lipoprotein (HDL) particles (pre- β HDL) with cholesterol. ABCG1 can load more mature HDL particles with cholesterol. Having taken up cholesterol through interaction with the ABC transporters in the artery wall, HDL particles can exit through the bloodstream, contributing to reverse cholesterol transport from lesional macrophages to the periphery. VLDL and LDL particles bearing ApoB can unload cholesterol from HDL particles through the action of cholesteryl ester transfer protein (CETP). Blockade of CETP can thus augment HDL levels, a process not yet known to produce clinical benefit. The ApoB-containing lipoproteins can promote clearance of cholesterol through capture by peripheral LDL receptors. Loss-of-function mutations in the enzyme PCSK9 (not shown) can increase the number of LDL receptors on peripheral cells, thereby augmenting the clearance of LDL.

Nicotinic acid raises HDL levels effectively and has shown some event reduction in clinical trials, but tolerability issues have limited its use. The recognition of G-protein-coupled receptors for nicotinic acid, and of β -hydroxybutyrate as an endogenous ligand of one such receptor, has not yet led to a therapeutic approach³⁷. Clinical end-point trials are testing whether adding nicotinic acid to standard care (usually including statin treatment) improves cardiovascular outcomes³⁸. One trial is using an extended-release preparation of nicotinic acid combined with a prostaglandin D_2 receptor antagonist, intended to reduce the cutaneous flushing that limits the acceptability of high-dose nicotinic acid for many patients³⁹.

Thus, despite considerable understanding of HDL and its metabolism, none of the pharmacological agents tested so far has offered a practical and proven way to reduce cardiovascular events. We must await the results of ongoing trials of approaches to raise HDL levels to reach this elusive goal.

Triglycerides on trial

Fasting or non-fasting triglyceride levels can predict cardiovascular events, but adjustment for other risk factors considerably weakens or even abolishes the association⁴⁰. Lifestyle changes such as weight loss, physical activity or low-carbohydrate diets can lower blood-circulating levels of triglycerides. The clinical benefits of

such lifestyle modifications probably result from a combination of mechanisms, so they cannot affirm triglycerides as a causal risk factor for atherosclerosis. Strict control of diabetes can also lessen hypertriglyceridaemia, yet tight glycaemic control may increase, rather than prevent, clinical complications of atherosclerosis in people with type 2 diabetes^{41,42}. Fibrates effectively decrease triglyceride levels, but trials of these drugs have proven disappointing in reducing clinical events, and have not shown reductions in mortality rates⁴³. Omega-3 fatty acids, prominent constituents of certain fish oils, reduce triglyceride levels and can limit cardiovascular events in some populations. These fatty acids also have anti-arrhythmic action, can mute inflammation and impair platelet aggregability — precluding a conclusion about the extent to which their clinical benefit arises from lowered triglyceride levels⁴⁴.

Recent evidence suggests that fractions of triglyceride-rich lipoproteins, particularly those that contain Apo-CIII, confer risk not conveyed by the total triglyceride level. Indeed, Apo-CIII acts as a pro-inflammatory mediator and an endogenous ligand of the Toll-like receptor 2 signalling pathway, which is implicated in the aggravation of mouse atherosclerosis^{45,46} (Fig. 2). In addition, very-low-density lipoprotein (VLDL) promotes transcription of the plasminogen activator-1 gene, leading to a reduced capacity to lyse thrombi⁴⁷.

Oxidation unproven to boost human atherogenesis

Experimental data indicate that oxidized LDL (oxLDL) within the arterial wall promotes the development of atherogenesis^{48–50} (Fig. 2). Yet direct evidence for the participation of LDL oxidation in human atherosclerosis remains scarce. Furthermore, the transition-metal-dependent chemistry that is commonly used to produce oxLDL in the laboratory may not reflect oxidative processes at work in the artery wall during atherogenesis. Such *in vitro* preparations of oxLDL vary by day, by donor and by laboratory, and constitute highly heterogeneous and undefined mixtures of biologically active substances.

Well-powered randomized clinical trials have consistently shown that antioxidant vitamins, such as vitamins E and C, do not reduce cardiovascular events over the timescale and doses studied^{51,52}. For many reasons, though, these studies of antioxidant vitamins do not themselves invalidate the oxidant hypothesis of atherosclerosis⁵³. One potent non-vitamin antioxidant (succinobucol) also failed to reduce coronary events in a large-scale trial⁵⁴.

Certain phospholipases may generate toxic or pro-inflammatory moieties from oxidized phospholipids associated with oxLDL. Clinical end-point trials with inhibitors of two of these enzymes are under way, but a small biomarker study with a lipoprotein-associated phospholipase A₂ inhibitor did not meet its prespecified coprimary end points⁵⁵ — again illustrating the quandary faced in pharmaceutical development about the design of smaller, short-term biomarker studies that can guide decisions about large clinical-outcome trials. Thus, despite the emergence of chemically defined active lipids and numerous publications about the potentially pro-atherogenic effects of oxLDL, its *in vivo* relevance in humans and its therapeutic manipulation remain speculative.

Translation from mice to humans

The discovery of T cells in human atheromata, and the subsequent identification of almost all cell types involved in innate and adaptive immunity in human plaques, raised the possibility that the immune system could participate in atherogenesis^{56–58} (Fig. 2). Markers of local adaptive and innate immune activation (such as major histocompatibility complex (MHC) antigens and leukocyte adhesion molecules) in plaques suggested the functional significance of immune cells in lesions. Studies in mice have established important modulatory roles of immunity in experimental atherosclerosis. T helper 1 (T_H1) cells, which produce the pro-inflammatory cytokines interferon- γ (IFN- γ) and TNF, have powerful pro-atherosclerotic actions in hypercholesterolaemic mice. This situation resembles that in several other inflammatory and autoimmune diseases, such as rheumatoid arthritis and type 1 diabetes, raising the question of whether specific antigens drive atherosclerosis. Two candidate autoantigens have emerged most prominently: LDL and heat-shock protein 60 (HSP60)⁵⁹. Cellular and humoral immune responses are mounted towards these antigens in humans and other experimental animals, and protective immunization strategies in mice have provided encouraging results^{60–62}. Of note, the cellular immune response to LDL specifically targets components of the native LDL particle rather than any oxidation-induced epitopes⁶³. Thus, fully oxidized LDL particles may not activate an adaptive immune response. For both LDL and HSP60, the extent to which the autoimmune response involves molecular mimicry with microbial pathogens remains undetermined.

Both stimulatory and inhibitory immune mechanisms operate during atherogenesis in hyperlipidaemic mice⁵⁷ (Fig. 2). Anti-inflammatory cytokines, such as IL-10 and transforming growth factor- β , counterbalance the pro-inflammatory pathways. Regulatory T (T_{reg}) cells produce anti-inflammatory signals that tend to counteract T_H1-cell production of IFN- γ — a cytokine that has long been implicated in atherogenesis^{64,65}. By contrast, some (but not all) studies indicate a pro-atherogenic role for T_H17 cells, which are a source of IL-17 (refs 66–70). B cells may have a protective role in atherosclerosis, as the removal of specific B-cell populations by splenectomy, for example, can aggravate atherosclerosis, and the transfer of antibodies reactive to LDL and its

components has shown atheroprotective effects in mice^{62,71}. Other experiments paradoxically show reduced atherosclerosis in mice after B-cell depletion⁷², suggesting that B-cell subpopulations may have contrasting roles in the pathogenesis of atherosclerosis.

Although these principles could apply to human atherosclerosis, direct translation from mouse studies is problematic. The accelerated atherogenesis in mice contrived to have hypercholesterolaemia requires cholesterol levels that far exceed those commonly encountered in the clinic, and does not reflect the chronic nature or complexity of the human disease⁷³. An approach to this problem would be to develop and adopt mouse models with lipoprotein metabolism closer to that in humans, or to use more moderate levels of dyslipidaemia over longer time periods. In studies of adaptive immunity, mice that are 'humanized' to carry human MHC genes may prove informative about antigen specificity.

The mouse immune system, although well understood and readily manipulated, diverges in many ways from that of humans⁷⁴. For example, humans lack the clear-cut T_H1 and T_H2 polarization found in mice. FOXP3 expression is a useful marker of T_{reg} cells in mice, but evidence suggests that it does not show the same degree of fidelity with T_{reg}-cell functionality in humans⁷⁵. The concept of polarization of macrophage functions probably does apply to humans, but the markers of classical (M1) versus alternative (M2) activation patterns in mice (for example, inducible nitric oxide synthase and arginase-1) differ from those in humans⁷⁶. Therefore, whenever possible, findings in mice should stimulate parallel studies in human biobanks and clinical studies. Identification of a putative disease-promoting molecule in human lesions, for example, or an increased risk associated with a genetic variant encoding the molecule, would lend value to a mechanistic study in gene-targeted mice.

Despite the disparities between the mouse and human immune systems (both innate and adaptive), clinical attempts at immunomodulation of atherosclerosis have begun. In particular, the recognition that humoral immunity can confer protection against experimental atherogenesis has spawned clinical trials involving the infusion of anti-LDL antibodies intended to reduce atherosclerosis. Vaccination studies with immunogens derived from LDL are also under way; experimental studies indicate that they evoke atheroprotective immunity involving cellular and humoral responses⁶².

Inflammation in atherosclerosis at the crossroads

A unifying view of the pathophysiology of atherosclerosis proposes that inflammation has a key role and transduces the effects of many known risk factors for the disease^{77,78}. Inflammatory signalling alters the behaviour of the intrinsic cells of the artery wall (endothelium and smooth muscle), and recruits further inflammatory cells that interact to promote lesion formation and complications (Figs 1 and 2). The application of biomarkers of inflammatory status, such as C-reactive protein (CRP), has lent clinical credence to this concept, but not without controversy. There remain many unanswered questions about the application of inflammation biology to human atherosclerosis. The association of inflammatory biomarkers with future risk of atherosclerotic complications does not demonstrate causality. Although the combined experimental and clinical evidence may convince some, the chicken-and-egg problem about causality remains unresolved.

Another open issue surrounds the use of anti-inflammatory therapy as a treatment for atherosclerosis. Many traditional anti-inflammatory therapies do not improve cardiovascular outcomes, and some may even aggravate atherosclerotic events. These observations, usually derived from post-hoc analyses of clinical studies or from mining observational databases, may reflect off-target actions of the agents studied — such as glucocorticoids, non-steroidal anti-inflammatory drugs, certain PPAR agonists or TNF inhibitors. 'Cardioprotective' doses of aspirin (50–150 mg daily) probably act as an antiplatelet agent rather than as a direct anti-inflammatory intervention. Despite the anti-inflammatory action of inhibitors of the cyclooxygenase-2 (COX-2) enzyme, the

BOX 2

Clinical trials evaluating anti-inflammatory agents

Here we describe two randomized, placebo-controlled trials that assess the efficacy of proven anti-inflammatory agents as cardiovascular therapeutic agents.

The cardiovascular inflammation reduction trial

This trial proposes to randomly allocate stable patients with post-myocardial infarction, who are receiving a complete standard-care regimen (including high-dose statin therapy), to either low-dose methotrexate (10–15 mg per week) or placebo⁹⁴. The treatment of rheumatoid arthritis routinely uses low-dose methotrexate, which has anti-inflammatory efficacy and an acceptable safety record among patients with similar age and co-morbidity status as individuals with stable coronary disease. Data from seven non-randomized observational cohorts of patients with rheumatoid arthritis or psoriatic arthritis demonstrate significant reductions in vascular event rates and cardiovascular death among individuals taking low-dose methotrexate rather than other disease-modifying agents. As low-dose methotrexate is a generic drug, a successful outcome for the trial would provide a simple, cost-effective method to address residual risk related to inflammation.

The Canakinumab Anti-inflammatory Thrombosis Outcomes Study

This study proposes to address directly whether, compared with placebo, IL-1 β inhibition can reduce the rates of recurrent myocardial infarction, stroke and cardiovascular-associated death among stable patients with coronary artery disease on a background of standard-care therapy (P.M.R., T. Thuren, A. Zalewski and P.L., manuscript in preparation). Canakinumab, a human monoclonal antibody, neutralizes the pro-inflammatory cytokine IL-1 β , which is implicated in atherothrombosis. Cholesterol crystals stimulate the NLRP3 inflammasome, which generates the active form of IL-1 β (refs 95, 96) (Fig. 1c). Canakinumab significantly reduces levels of inflammatory biomarkers such as CRP, and is currently used to treat inherited IL-1 β -driven inflammatory diseases such as Muckle-Wells syndrome. Because IL-1 β may participate in autoimmune processes related to pancreatic dysfunction and insulin resistance, this study also has a secondary prespecified end point of new-onset diabetes. If successful, the trial would support the inflammatory hypothesis of atherothrombosis, and provide a new cytokine-based therapy for the secondary prevention of cardiovascular disease and new-onset diabetes.

pro-thrombotic effect of inhibiting prostacyclin production may contribute to increased cardiovascular morbidity⁷⁹.

Statins effectively lower LDL and CRP levels in humans. Analyses of several large studies of statins in primary- and secondary-prevention populations suggest that some of their clinical benefit accrues from an anti-inflammatory action distinct from LDL lowering^{80,81}. The hypothesis that an anti-inflammatory intervention can reduce cardiovascular events independent of lipoprotein effects still requires rigorous testing. Thus, despite hundreds of studies affirming a role for inflammation in atherosclerosis in mice, and many intriguing observations in humans, Koch's postulates remain unfulfilled.

Ultimately, testing the inflammatory hypothesis of atherothrombosis will require a series of randomized, placebo-controlled trials that evaluate proven anti-inflammatory agents without confounding effects on cholesterol or platelet function as cardiovascular therapeutic agents. At least two such trials (Box 2) should begin soon, targeting a high-risk population with persistent inflammation, thus limiting the intervention to those most likely to benefit.

Animal experiments versus human disease

What lessons can we learn from the frustrations in clinical application of advances in atherosclerosis biology, and how can we tighten the coupling between scientific advances and clinical practice? Animal experiments have proven indispensable to studies of disease mechanisms, but we must not forget their limitations. Too often, the pharmaceutical or biotechnology sector adopts or abandons targets or strategies on the basis of uncritical acceptance of the results of animal studies. The recognition of animal preparations as 'models' of human disease requires considerable scepticism. For example, atherosclerotic lesions in the commonly used genetically modified mice seldom develop plaque disruption with thrombosis — a mechanism that commonly complicates the human disease. Mouse studies generally focus on the aorta and proximal great vessels, whereas the most important clinical consequences of atherosclerosis in humans arise from lesions in the coronary, carotid and cerebral arteries. The structure and hydrodynamics of these smaller muscular arteries, and even the embryonic origin of their SMCs, differ markedly from the large elastic arteries usually analysed in mouse studies. The proximal left anterior descending coronary artery in humans, a frequent site of lesion formation, characteristically contains a considerable population of intimal SMCs, even in early life

— a major difference from mouse arteries. Coronary arterial SMCs arise from the proepicardial organ, not from the splanchnic mesoderm as do those in the descending aorta⁵. In contrast to the aorta, in which flow predominates in systole, flow in the coronary arteries occurs mainly in diastole, and the heart's arteries experience compressive forces during systole. Human coronary arteries usually lie in an extensive pedicle of perivascular fat that may provoke outside-in signalling.

These distinctions by no means indicate that we should discard animal studies of atherosclerosis, or forgo the immense power of mouse genetics to pose questions about pathophysiology. But animal studies do require judicious interpretation, and recognition of their limitations, when extrapolating to human disease. Experimental atherosclerosis in animals allows the rigorous testing of mechanistic hypotheses, but does not mimic the human condition entirely.

The importance of biomarkers

Advances in proteomic, metabolomic and genetic technologies have led to the accelerated identification of putative biomarkers of disease and risk factors for complications, and the targeting or improved efficacy of therapies. Selective harnessing of biomarkers can help to gauge the relevance of experimental results to human disease. For example, a highly sensitive assay to measure CRP has helped to translate to the clinic the results of decades of laboratory studies that implicated inflammatory pathways in the pathogenesis of atherosclerosis.

Genome-wide association studies have reproducibly identified and validated regions of the human genome that associate with the risk of myocardial infarction. For example, the chromosome 9p21 region, which consistently associates with a greater risk of myocardial infarction, has begun to yield new biological insight⁸², as have several variants associated with lipoprotein disorders^{83,84}. This unbiased approach will identify therapeutic targets that have eluded the classical model of drug development. Most identified genetic risk factors contribute moderately to disease and do not yet justify population screening^{85,86}. Because many risk-conferring genes may make only small contributions to risk, common variants (compared with rarer mutations) may not prove useful in risk prediction.

In addition to bridging laboratory and human studies, the application of biomarkers could help to advance the treatment of people with, or at risk of, atherosclerosis by improving prognostication, by assessing the need for and intensity of therapy, by individualizing the use of specific

therapies, and by helping to develop new therapeutics. For example, including CRP with conventional risk factors improves risk prediction for atherosclerotic events, both in people with and without established disease. Sensitive cardiac troponin measurements can detect levels of ischaemic damage far beneath the clinical threshold and convey incremental risk information. The serum protein cystatin-C and brain natriuretic peptides may also have clinical use in risk prediction.

A randomized, multicentre trial that finished in 2008 illustrates how the inflammation biomarker CRP can identify individuals who are not eligible for therapy according to traditional approaches, but who could benefit from treatment with a statin that has potent LDL-lowering and anti-inflammatory effects⁸⁷. Yet this study could not, nor was it designed to, determine the mechanism of event reduction, reinforcing the need for future trials of anti-inflammatory agents that do not alter lipid levels.

Beyond risk prediction and targeting of therapy, the application of biomarkers may help the field of cardiovascular therapeutics to confront the enormous challenge it faces — to discover and develop therapeutics for modulating atherosclerosis. Owing to the success of the current standard of care, clinical end-point trials now generally involve patient populations with lower event rates. Consequently, clinical trials that pit new strategies for reducing atherosclerotic events against the current standard of care will require greater numbers of participants and longer study durations than in previous eras, with attendant greater expense. Better validation methods are needed for the targets arising from the burgeoning basic science of atherosclerosis in humans. Genome-wide association studies are identifying numerous new targets for drug development. Application of this knowledge in humans will require methods to determine whether interventions will affect their intended targets, to optimize doses, and to obtain early signals compatible with clinical benefit in pilot studies of fewer subjects and shorter duration. Such methods could inform decisions about which agents should move forward into increasingly expensive and arduous large clinical trials.

There is no single optimum biomarker for reporting the possible clinical efficacy of new therapeutics. Biomarker selection for these purposes should reflect the mechanisms under scrutiny. Atheroma volume measured by intravascular ultrasound, for example, might be an appropriate biomarker for an intervention designed to unload lipid from plaques, such as an Apo-AI mimetic; the level of lipoprotein-associated phospholipase A₂ in peripheral blood leukocytes would be an appropriate biomarker for determining the dose range of an inhibitor of that enzyme (in the absence of direct access to the relevant tissue — the human plaque itself); and CRP measurement could serve as a marker to assess an anti-inflammatory intervention that may not affect plaque size.

Biomarkers include traditional analytes in body fluids and anthropometric measurements, and can include, by some definitions, structural variables measured by imaging. Biomarkers are unlikely to provide surrogate end points of efficacy that prove acceptable to regulatory agencies for the registration of new therapeutics in the foreseeable future, but they should assist in bridging the translational gap.

In addition to anatomical imaging, harnessing biological processes to provide imaging targets (molecular imaging) may help to test mechanistic hypotheses in humans, and provide early signals about the efficacy of interventions in small and short pilot studies⁸⁸. With respect to atherosclerosis, molecular imaging using different platforms has proven promising in visualizing adhesion molecules, integrins, phagocytosis, proteases, reactive oxygen species and modified lipoproteins⁸⁸. The modalities that have shown potential in this regard include isotope-tagged ligands, paramagnetic agents visualized by magnetic resonance imaging, contrast-enhanced ultrasound and near-infrared fluorescent probes. Such methodologies might result in crucial information for phase II drug development, including ascertainment of *in vivo* targeting in humans (not just in the blood, but in the atheroma itself), and provide human data about doses for clinical end-point studies.

Although molecular imaging of atherosclerosis shows promise in animals, it faces great hurdles to clinical translation. The production

of molecular probes for human use requires good manufacturing processes, toxicology evaluation and, often, the extension of innovative imaging platforms beyond the pilot stage. Overcoming these barriers requires resources beyond the reach of most academic groups, necessitating governmental, industrial or philanthropic support.

Clinical trials as a laboratory for discovery

Clinical trials should be used more often as an early scientific probe, not just as a pathway to the commercialization of pharmaceuticals or for evaluating comparative efficacy of established agents. Although daunting to design, fund and conduct, clinical trials constitute the ultimate translational tool. The publications reporting many laboratory studies convey an optimistic speculation about clinical extrapolation. A deep and wide chasm separates the promises in these sentences and a randomized, prospective clinical trial that tests the conjecture. Prohibitive practical limitations impose themselves, and not many hypotheses arising from laboratory studies will undergo such rigorous clinical evaluation; hence, it is necessary to harness biomarkers more effectively to identify strategies that have the most promise for clinical translation. Clinical trialists should strive to archive biobanks and build biomarker sub-studies into clinical trials whenever possible, to allow post-hoc data mining, generate new hypotheses, and test those mechanistic hypotheses already specified.

The increasing expense of clinical end-point trials, driven by the considerations explained above, constitutes a major limitation to the translation of biological advances to atherosclerosis treatment. The daunting costs of cardiovascular clinical trials have diverted investments of the pharmaceutical industry to other therapeutic areas, reducing the discovery effort and limiting the number of approaches that will undergo clinical evaluation. Models for public support of trials to test crucial hypotheses, including those that may have little commercial appeal, for funding of ancillary mechanistic studies or sub-studies, and for improvements in trial designs to render them less costly would help to surmount these barriers.

The biological insights and experimental progress in understanding the mechanisms of atherosclerosis and its complications have advanced markedly. But full understanding of the applicability of laboratory findings to humans and the realization of therapeutic promise require another investigative dimension. We must reach beyond the tools available in the laboratory to probe pathophysiology, and more urgently strive to bridge the gap to human disease. ■

1. Glass, C. K. & Witztum, J. L. Atherosclerosis: the road ahead. *Cell* **104**, 503–516 (2001).
2. Lusis, A. J. Atherosclerosis. *Nature* **407**, 233–241 (2000).
3. Libby, P. Vascular biology of atherosclerosis: overview and state of the art. *Am. J. Cardiol.* **91**, 3–6 (2003).
4. Tabas, I., Williams, K. J. & Boren, J. Subendothelial lipoprotein retention as the initiating process in atherosclerosis: update and therapeutic implications. *Circulation* **116**, 1832–1844 (2007).
5. Majesky, M. W. Developmental basis of vascular smooth muscle diversity. *Arterioscler. Thromb. Vasc. Biol.* **27**, 1248–1258 (2007).
6. Gimbrone, M. A. Jr, Topper, J. N., Nagel, T., Anderson, K. R. & Garcia-Cardena, G. Endothelial dysfunction, hemodynamic forces, and atherogenesis. *Ann. NY Acad. Sci.* **902**, 230–240 (2000).
7. Swirski, F. K. *et al.* Ly-6C^{hi} monocytes dominate hypercholesterolemia-associated monocytosis and give rise to macrophages in atheromata. *J. Clin. Invest.* **117**, 195–205 (2007).
8. Tacke, F. *et al.* Monocyte subsets differentially employ CCR2, CCR5, and CX3CR1 to accumulate within atherosclerotic plaques. *J. Clin. Invest.* **117**, 185–194 (2007).
9. Bouhlel, M. A. *et al.* PPAR γ activation primes human monocytes into alternative M2 macrophages with anti-inflammatory properties. *Cell Metab.* **6**, 137–143 (2007).
10. Tabas, I. Macrophage death and defective inflammation resolution in atherosclerosis. *Nature Rev. Immunol.* **10**, 36–46 (2010).
11. Libby, P. Molecular and cellular mechanisms of the thrombotic complication of atherosclerosis. *J. Lipid Res.* **50**, S352–S357 (2009).
12. Steinberg, D. *The Cholesterol Wars: the Skeptics vs. the Preponderance of Evidence* 1st edn (Elsevier, 2007).
13. Goldstein, J. L. & Brown, M. S. The LDL receptor. *Arterioscler. Thromb. Vasc. Biol.* **29**, 431–438 (2009).
14. Brown, M. S. & Goldstein, J. L. Heart attacks: gone with the century? *Science* **272**, 629 (1996).

15. Libby, P. The forgotten majority: unfinished business in cardiovascular risk reduction. *J. Am. Coll. Cardiol.* **46**, 1225–1228 (2005).
16. Cannon, C. P. *et al.* Intensive versus moderate lipid lowering with statins after acute coronary syndromes. *N. Engl. J. Med.* **350**, 1495–1504 (2004).
This study demonstrates the residual risk for patients who have survived an acute coronary syndrome despite intensive statin treatment.
17. Nissen, S. E. *et al.* Effect of very high-intensity statin therapy on regression of coronary atherosclerosis: the ASTEROID trial. *J. Am. Med. Assoc.* **295**, 1556–1565 (2006).
18. Tall, A. R., Yvan-Charvet, L., Terasaka, N., Pagler, T. & Wang, N. HDL, ABC transporters, and cholesterol efflux: implications for the treatment of atherosclerosis. *Cell Metab.* **7**, 365–375 (2008).
19. Rye, K. A., Bursill, C. A., Lambert, G., Tabet, F. & Barter, P. J. The metabolism and anti-atherogenic properties of HDL. *J. Lipid Res.* **50**, S195–S200 (2009).
20. Vaisar, T. *et al.* Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL. *J. Clin. Invest.* **117**, 746–756 (2007).
21. Brewer, H. B. Jr. High-density lipoproteins: a new potential therapeutic target for the prevention of cardiovascular disease. *Arterioscler. Thromb. Vasc. Biol.* **24**, 387–391 (2004).
22. Chapman, M. J., Le Goff, W., Guerin, M. & Kontush, A. Cholesteryl ester transfer protein: at the heart of the action of lipid-modulating therapy with statins, fibrates, niacin, and cholesteryl ester transfer protein inhibitors. *Eur. Heart J.* **31**, 149–164 (2010).
23. Jahangiri, A. *et al.* HDL remodeling during the acute phase response. *Arterioscler. Thromb. Vasc. Biol.* **29**, 261–267 (2009).
24. Asztalos, B. F. *et al.* Differential effects of HDL subpopulations on cellular ABCA1- and SR-BI-mediated cholesterol efflux. *J. Lipid Res.* **46**, 2246–2253 (2005).
25. Barter, P. J. *et al.* Effects of torcetrapib in patients at high risk for coronary events. *N. Engl. J. Med.* **357**, 2109–2122 (2007).
26. Barter, P. J. & Kastelein, J. J. Targeting cholesteryl ester transfer protein for the prevention and management of cardiovascular disease. *J. Am. Coll. Cardiol.* **47**, 492–499 (2006).
27. Cannon, C. P. *et al.* Safety of anacetrapib in patients with or at high risk for coronary heart disease. *N. Engl. J. Med.* **363**, 2406–2415 (2010).
28. Nissen, S. E. *et al.* Effect of recombinant ApoA-I Milano on coronary atherosclerosis in patients with acute coronary syndromes: a randomized controlled trial. *J. Am. Med. Assoc.* **290**, 2292–2300 (2003).
29. Tardif, J. C. *et al.* Effects of reconstituted high-density lipoprotein infusions on coronary atherosclerosis: a randomized controlled trial. *J. Am. Med. Assoc.* **297**, 1675–1682 (2007).
30. Navab, M. *et al.* Human apolipoprotein A1 mimetic peptides for the treatment of atherosclerosis. *Curr. Opin. Investig. Drugs* **4**, 1100–1104 (2003).
31. Staels, B. Fibrates in CVD: a step towards personalised medicine. *Lancet* **375**, 1847–1848 (2010).
32. Frick, M. H. *et al.* Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged men with dyslipidemia. *N. Engl. J. Med.* **317**, 1237–1245 (1987).
33. Robins, S. J. *et al.* Relation of gemfibrozil treatment and lipid levels with major coronary events. VA-HIT: a randomized controlled trial. *J. Am. Med. Assoc.* **285**, 1585–1591 (2001).
34. Jones, P. H. & Davidson, M. H. Reporting rate of rhabdomyolysis with fenofibrate + statin versus gemfibrozil + any statin. *Am. J. Cardiol.* **95**, 120–122 (2005).
35. Keech, A. *et al.* Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. *Lancet* **366**, 1849–1861 (2005).
36. Ginsberg, H. N. *et al.* Effects of combination lipid therapy in type 2 diabetes mellitus. *N. Engl. J. Med.* **362**, 1563–1574 (2010).
37. Taggart, A. K. *et al.* (D)- β -Hydroxybutyrate inhibits adipocyte lipolysis via the nicotinic acid receptor PUMA-G. *J. Biol. Chem.* **280**, 26649–26652 (2005).
This paper identifies a target of nicotinic acid action that may provide mechanistic insight into its mode of action.
38. Brown, B. G. & Zhao, X. Q. Nicotinic acid, alone and in combinations, for reduction of cardiovascular risk. *Am. J. Cardiol.* **101**, 58B–62B (2008).
39. Duffy, D. & Rader, D. J. Update on strategies to increase HDL quantity and function. *Nature Rev. Cardiol.* **6**, 455–463 (2009).
40. Sarwar, N. *et al.* Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. *Circulation* **115**, 450–458 (2007).
41. Gerstein, H. C. *et al.* Effects of intensive glucose lowering in type 2 diabetes. *N. Engl. J. Med.* **358**, 2545–2559 (2008).
42. Duckworth, W. *et al.* Glucose control and vascular complications in veterans with type 2 diabetes. *N. Engl. J. Med.* **360**, 129–139 (2009).
43. Jun, M. *et al.* Effects of fibrates on cardiovascular outcomes: a systematic review and meta-analysis. *Lancet* **375**, 1875–1884 (2010).
44. Siscovick, D. S., Lemaitre, R. N. & Mozaffarian, D. The fish story: a diet-heart hypothesis with clinical implications: n-3 polyunsaturated fatty acids, myocardial vulnerability, and sudden death. *Circulation* **107**, 2632–2634 (2003).
45. Kawakami, A. *et al.* Toll-like receptor 2 mediates apolipoprotein CIII-induced monocyte activation. *Circ. Res.* **103**, 1402–1409 (2008).
46. Mullick, A. E., Tobias, P. S. & Curtiss, L. K. Modulation of atherosclerosis in mice by Toll-like receptor 2. *J. Clin. Invest.* **115**, 3149–3156 (2005).
47. Eriksson, P., Nilsson, L., Karpe, F. & Hamsten, A. Very-low-density lipoprotein response element in the promoter region of the human plasminogen activator inhibitor-1 gene implicated in the impaired fibrinolysis of hypertriglyceridemia. *Arterioscler. Thromb. Vasc. Biol.* **18**, 20–26 (1998).
48. Berliner, J. A. & Watson, A. D. A role for oxidized phospholipids in atherosclerosis. *N. Engl. J. Med.* **353**, 9–11 (2005).
49. Steinberg, D. The LDL modification hypothesis of atherogenesis: an update. *J. Lipid Res.* **50**, S376–S381 (2009).
A balanced weighing of the oxidized-lipid hypothesis of atherosclerosis from a pioneer in the field.
50. Steinberg, D. & Witztum, J. L. Oxidized low-density lipoprotein and atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **30**, 2311–2316 (2010).
51. Lonn, E. *et al.* Effects of long-term vitamin E supplementation on cardiovascular events and cancer: a randomized controlled trial. *J. Am. Med. Assoc.* **293**, 1338–1347 (2005).
52. Lonn, E. *et al.* Homocysteine lowering with folic acid and B vitamins in vascular disease. *N. Engl. J. Med.* **354**, 1567–1577 (2006).
53. Steinberg, D. & Witztum, J. L. Is the oxidative modification hypothesis relevant to human atherosclerosis? *Circulation* **105**, 2107–2111 (2002).
54. Tardif, J. C. *et al.* Effects of succinobucol (AGI-1067) after an acute coronary syndrome: a randomised, double-blind, placebo-controlled trial. *Lancet* **371**, 1761–1768 (2008).
55. Serruys, P. W. *et al.* Effects of the direct lipoprotein-associated phospholipase A₂ inhibitor darapladib on human coronary atherosclerotic plaque. *Circulation* **118**, 1172–1182 (2008).
56. Jonasson, L., Holm, J., Skalli, O., Bondjers, G. & Hansson, G. K. Regional accumulations of T cells, macrophages, and smooth muscle cells in the human atherosclerotic plaque. *Arteriosclerosis* **6**, 131–138 (1986).
57. Hansson, G. K. & Libby, P. The immune response in atherosclerosis: a double-edged sword. *Nature Rev. Immunol.* **6**, 508–519 (2006).
58. Hartvigsen, K. *et al.* The role of innate immunity in atherogenesis. *J. Lipid Res.* **50**, S388–S393 (2008).
59. Andersson, J., Libby, P. & Hansson, G. K. Adaptive immunity and atherosclerosis. *Clin. Immunol.* **134**, 33–46 (2010).
60. Palinski, W., Miller, E. & Witztum, J. L. Immunization of low density lipoprotein (LDL) receptor-deficient rabbits with homologous malondialdehyde-modified LDL reduces atherogenesis. *Proc. Natl Acad. Sci. USA* **92**, 821–825 (1995).
61. Maron, R. *et al.* Mucosal administration of heat shock protein-65 decreases atherosclerosis and inflammation in aortic arch of low-density lipoprotein receptor-deficient mice. *Circulation* **106**, 1708–1715 (2002).
62. Hansson, G. K. & Nilsson, J. Vaccination against atherosclerosis? Induction of atheroprotective immunity. *Semin. Immunopathol.* **31**, 95–101 (2009).
63. Hermansson, A. *et al.* Inhibition of T cell response to native low-density lipoprotein reduces atherosclerosis. *J. Exp. Med.* **207**, 1081–1093 (2010).
64. Robertson, A. K. *et al.* Disruption of TGF- β signaling in T cells accelerates atherosclerosis. *J. Clin. Invest.* **112**, 1342–1350 (2003).
65. Ait-Oufella, H. *et al.* Natural regulatory T cells control the development of atherosclerosis in mice. *Nature Med.* **12**, 178–180 (2006).
66. van Es, T. *et al.* Attenuated atherosclerosis upon IL-17R signaling disruption in LDLr deficient mice. *Biochem. Biophys. Res. Commun.* **388**, 261–265 (2009).
67. Taleb, S., Tedgui, A. & Mallat, Z. Interleukin-17: friend or foe in atherosclerosis? *Curr. Opin. Lipidol.* **21**, 404–408 (2010).
68. Taleb, S. *et al.* Loss of SOCS3 expression in T cells reveals a regulatory role for interleukin-17 in atherosclerosis. *J. Exp. Med.* **206**, 2067–2077 (2009).
69. Madhur, M. S. *et al.* Role of interleukin 17 in inflammation, atherosclerosis, and vascular function in apolipoprotein E-deficient mice. *Arterioscler. Thromb. Vasc. Biol.* doi:10.1161/ATVBAHA.111.227629 2011 (7 April 2011).
70. Cheng, X. *et al.* Inhibition of IL-17A in atherosclerosis. *Atherosclerosis* **215**, 471–474 (2011).
71. Caligiuri, G., Nicoletti, A., Poirier, B. & Hansson, G. K. Protective immunity against atherosclerosis carried by B cells of hypercholesterolemic mice. *J. Clin. Invest.* **109**, 745–753 (2002).
72. Ait-Oufella, H. *et al.* B cell depletion reduces the development of atherosclerosis in mice. *J. Exp. Med.* **207**, 1579–1587 (2010).
73. Bentzon, J. F. & Falk, E. Atherosclerotic lesions in mouse and man: is it the same disease? *Curr. Opin. Lipidol.* **21**, 434–440 (2010).
74. Nussenblatt, R. B. *et al.* National Institutes of Health Center for Human Immunology Conference, September 2009. *Ann. NY Acad. Sci.* **1200**, E1–E23 (2010).
A recent compilation of some of the distinctions between the human and mouse immune systems.
75. Tran, D. Q., Ramsey, H. & Shevach, E. M. Induction of FOXP3 expression in naive human CD4⁺FOXP3⁺ T cells by T-cell receptor stimulation is transforming growth factor- β -dependent but does not confer a regulatory phenotype. *Blood* **110**, 2983–2990 (2007).
76. Raes, G., Van den Bergh, R., De Baetselier, P. & Ghassabeh, G. H. Arginase-1 and Ym1 are markers for murine, but not human, alternatively activated myeloid cells. *J. Immunol.* **174**, 6561–6562 (2005).
77. Hansson, G. K. Inflammation, atherosclerosis, and coronary artery disease. *N. Engl. J. Med.* **352**, 1685–1695 (2005).
78. Libby, P. & Ridker, P. M. Inflammation and atherothrombosis. *J. Am. Coll. Cardiol.* **48**, 33–46 (2006).
79. Grosser, T., Fries, S. & FitzGerald, G. A. Biological basis for the cardiovascular consequences of COX-2 inhibition: therapeutic challenges and opportunities. *J. Clin. Invest.* **116**, 4–15 (2006).
80. Ridker, P. M. *et al.* C-reactive protein levels and outcomes after statin therapy. *N. Engl. J. Med.* **352**, 20–28 (2005).
81. Ridker, P. M. *et al.* Reduction in C-reactive protein and LDL cholesterol and cardiovascular event rates after initiation of rosuvastatin: a prospective study of the JUPITER trial. *Lancet* **373**, 1175–1182 (2009).
82. Harismendy, O. *et al.* 9p21 DNA variants associated with coronary artery disease impair interferon- γ signalling response. *Nature* **470**, 264–268 (2011).

83. Schunkert, H. *et al.* Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* **117**, 1675–1684 (2008).
84. Kathiresan, S. *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature Genet.* **40**, 189–197 (2008).
85. Morgan, T. M., Krumholz, H. M., Lifton, R. P. & Spertus, J. A. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *J. Am. Med. Assoc.* **297**, 1551–1561 (2007).
This paper describes the lack of reproducibility of associations between single-nucleotide polymorphisms and atherosclerotic outcomes.
86. Paynter, N. P. *et al.* Association between a literature-based genetic risk score and cardiovascular events in women. *J. Am. Med. Assoc.* **303**, 631–637 (2010).
This study shows that a panel of reproducible genetic variants from genome-wide association studies does not improve cardiovascular risk prediction.
87. Ridker, P. M. *et al.* Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N. Engl. J. Med.* **359**, 2195–2207 (2008).
88. Libby, P., Di Carli, M. F. & Weissleder, R. The vascular biology of atherosclerosis and imaging targets. *J. Nucl. Med.* **51** (suppl. 1), 33S–37S (2010).
89. Altmann, S. W. *et al.* Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol absorption. *Science* **303**, 1201–1204 (2004).
90. Cohen, J. C., Boerwinkle, E., Mosley, T. H. Jr & Hobbs, H. H. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* **354**, 1264–1272 (2006).
This study identifies a new target for LDL lowering and provides evidence that life-long low LDL levels confer considerable cardiovascular protection, reinforcing and extending the results of short-term drug-intervention trials.
91. Brown, M. S. & Goldstein, J. L. Lowering LDL — not only how low, but how long? *Science* **311**, 1721–1723 (2006).
92. Wang, C. Y., Liu, P. Y. & Liao, J. K. Pleiotropic effects of statin therapy: molecular mechanisms and clinical results. *Trends Mol. Med.* **14**, 37–44 (2008).
93. Parmar, K. M. *et al.* Integration of flow-dependent endothelial phenotypes by Kruppel-like factor 2. *J. Clin. Invest.* **116**, 49–58 (2006).
94. Ridker, P. M. Testing the inflammatory hypothesis of atherothrombosis: scientific rationale for the cardiovascular inflammation reduction trial (CIRT). *J. Thromb. Haemost.* **7** (suppl. s1), 332–339 (2009).
95. Duwell, P. *et al.* NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature* **464**, 1357–1361 (2010).
96. Rajamaki, K. *et al.* Cholesterol crystals activate the NLRP3 inflammasome in human macrophages: a novel link between cholesterol metabolism and inflammation. *PLoS ONE* **5**, e11765 (2010).

Acknowledgements We thank S. Karwacki for editorial assistance.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to P.L. (plibby@rics.bwh.harvard.edu).

Heart regeneration

Michael A. Laflamme¹ & Charles E. Murry^{1,2,3}

Heart failure plagues industrialized nations, killing more people than any other disease. It usually results from a deficiency of specialized cardiac muscle cells known as cardiomyocytes, and a robust therapy to regenerate lost myocardium could help millions of patients every year. Heart regeneration is well documented in amphibia and fish and in developing mammals. After birth, however, human heart regeneration becomes limited to very slow cardiomyocyte replacement. Several experimental strategies to remuscularize the injured heart using adult stem cells and pluripotent stem cells, cellular reprogramming and tissue engineering are in progress. Although many challenges remain, these interventions may eventually lead to better approaches to treat or prevent heart failure.

Heart regeneration has been intensely investigated, and extremely controversial, for more than 150 years¹. In pursuit of this subject, the heart has been stabbed, snipped, contused, cauterized, coagulated, frozen, injected with toxins, infected and infarcted, in species ranging from marine invertebrates to horses^{2,3}. Why has this proven to be such a difficult challenge? The heart is one of the least regenerative organs in the body, so if there is a regenerative response, it is small in comparison to that seen in many other tissues, such as liver, skeletal muscle, lung, gut, bladder, bone or skin. For most investigators, the question is about whether there is no regeneration, which is intrinsically difficult to prove, or whether it occurs but at very low rates, which is not easy to detect but possible using highly sensitive approaches.

This is more than an academic argument. Heart failure is a burgeoning public health problem, and some predict that it will reach epidemic proportions as our population ages. Cardiomyocyte deficiency underlies most causes of heart failure. The human left ventricle has 2–4 billion cardiomyocytes, and a myocardial infarction can wipe out 25% of these in a few hours⁴. Disorders of cardiac overload such as hypertension or valvular heart disease kill cardiomyocytes slowly over many years⁵, and ageing is associated with the loss of ~1 g of myocardium (about 20 million cardiomyocytes) per year in the absence of specific heart disease⁶. If the human heart has even a small innate regenerative response, it may be possible to exploit this therapeutically to enhance the heart's function. This fundamental motivation has kept investigators pursuing rare events for more than a century.

Over the past 15 years, researchers have taken a more interventional approach to the injured heart, creating the field of cardiac repair. The ultimate goal of cardiac repair is to regenerate the myocardium after injury to prevent or treat heart failure. This interdisciplinary field draws from advances in areas such as stem cells, developmental biology and biomaterials in an attempt to create new myocardium that is electrically and mechanically integrated into the heart. Cardiac repair has moved rapidly from studies in experimental animals to clinical trials involving thousands of patients. In this Review, we summarize the evidence for heart regeneration in animal models and humans. We discuss the status of research using adult stem cells and pluripotent stem cells for cardiac repair in experimental animals, and explore the promises and problems of cellular reprogramming and tissue engineering. Clinical trials will be covered only briefly, owing to space limitations, so we refer interested readers to recent reviews on this topic^{7,8}.

Heart regeneration in amphibia and fish

Unlike humans, many amphibia and fish readily regenerate limbs, appendages and internal organs after injury. There is a long history of research on amphibian heart regeneration⁹; more recently, the zebrafish has proven to be a particularly useful model, given its substantial regenerative capacity and amenability to genetic manipulation¹⁰. The zebrafish heart fully regenerates after the surgical amputation of the cardiac apex — an injury that corresponds to a loss of approximately 20% of the total ventricular mass¹⁰. In the low-pressure zebrafish heart, this large wound is effectively sealed by an initial fibrin clot, which is gradually replaced by *de novo* regenerated heart tissue rather than by scar tissue^{10–13}.

Not surprisingly, this regenerative response involves a substantial amount of cardiomyocyte proliferation. Even at baseline levels, zebrafish cardiomyocytes show a much higher degree of cell-cycle activity than equivalent cells from their mammalian counterparts. A recent study showed that approximately 3% of cardiomyocytes in the compact myocardium of uninjured adult zebrafish hearts incorporate the thymidine analogue bromodeoxyuridine (BrdU) during a seven-day pulse-labelling experiment. Two weeks after amputation of the cardiac apex, the fraction of BrdU-positive cardiomyocytes had increased by tenfold, and this parameter remained as high as 20% as late as one month after injury¹⁰.

Initial experiments suggested that undifferentiated progenitor cells were the principal source of regenerating cardiomyocytes in zebrafish¹¹, but two recent genetic fate-mapping studies unambiguously demonstrated that pre-existing committed cardiomyocytes are instead the main source^{12,13} (Box 1). The two groups independently generated transgenic zebrafish in which the cardiomyocyte-specific *cmlc2* (also known as *myl7*) promoter drives the expression of tamoxifen-inducible Cre recombinase. These animals were crossed with a reporter line, in which Cre-mediated excision of a *loxP*-flanked stop sequence induces constitutive expression of green fluorescent protein (GFP). In the offspring of this cross, all pre-existing cardiomyocytes and their progeny can be induced to express GFP by tamoxifen treatment. If the regenerated myocardium were derived from undifferentiated progenitor cells, the new ventricular apex should be GFP⁺. Instead, both groups found that the vast majority of the newly regenerated cardiomyocytes were GFP⁺ (refs 12, 13). Thus, heart regeneration in zebrafish is principally mediated by the proliferation of pre-existing cardiomyocytes, rather than the generation of new cardiomyocytes from stem cells.

¹Department of Pathology, Center for Cardiovascular Biology, Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington 98109, USA. ²Department of Bioengineering, Center for Cardiovascular Biology, Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington 98109, USA. ³Department of Medicine/Cardiology, Center for Cardiovascular Biology, Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington 98109, USA.

BOX 1

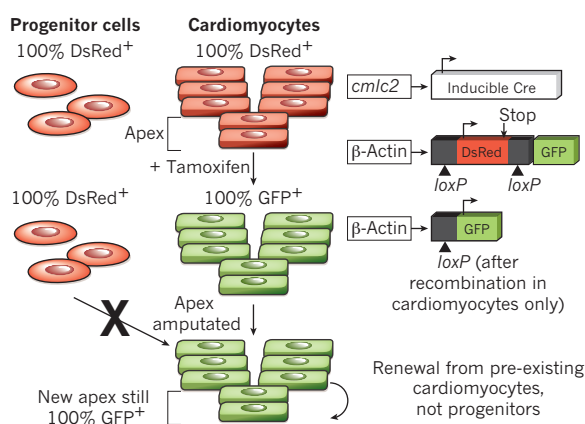
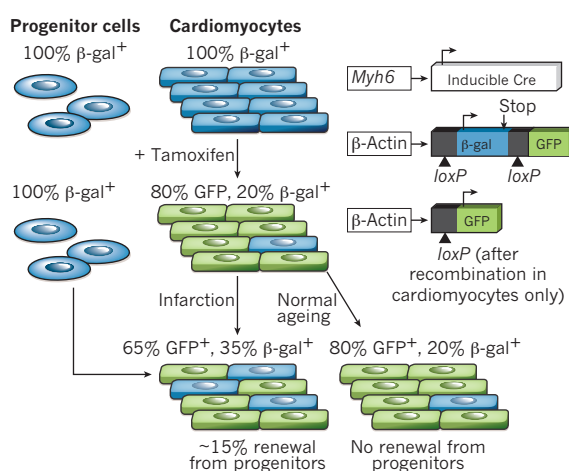
Genetic fate mapping in heart regeneration

Genetic fate mapping has proven to be an invaluable tool for dissecting the mechanisms of endogenous cardiac repair in model organisms, and several laboratories have used an elegant strategy based on the conditional Cre-*loxP* system, which allows both temporal and cell-type-specific control of reporter expression.

After amputation, the apex of the zebrafish heart can fully regenerate. To determine the source of the newly proliferating cardiomyocytes that underlies this regeneration, a zebrafish strain carrying two transgenes was created (see Figure, **a**). In one transgene, the cardiomyocyte-specific *cmlc2* promoter drives the expression of tamoxifen-inducible Cre recombinase. In the second transgene, the constitutive β -actin promoter initially drives expression of the red fluorescent DsRed protein. Cre recombinase induces the excision of *loxP*-flanked stop sequences, causing a permanent switch from constitutive DsRed to constitutive green fluorescent protein (GFP) expression¹². Thus, when the transgenic zebrafish was pulsed with tamoxifen, all of its cardiomyocytes and their descendants expressed GFP. By contrast, cardiomyogenic progenitor cells should remain DsRed⁺GFP⁻, because the cardiomyocyte-specific *cmlc2* promoter would not be active in these undifferentiated cells. If progenitor cells later contributed to cardiomyocyte renewal after injury, one would expect those cardiomyocytes to also be DsRed⁺GFP⁻. Instead, after amputation of the apex, the new apical myocardium was 100% GFP⁺, indicating that heart regeneration in the zebrafish results from the expansion of pre-existing cardiomyocytes, not from the recruitment of cardiomyogenic precursors. Another independent group reached the same conclusions using a similar experimental design¹³.

An analogous genetic fate-mapping approach was used to investigate the mechanisms of cardiac regeneration in mammalian hearts²². Here, a double-transgenic mouse was generated in which the cardiomyocyte-specific *Myh6* promoter drives tamoxifen-inducible Cre recombinase, and Cre-mediated excision of *loxP*-flanked stop sequences induces a switch from constitutive β -galactosidase (β -gal) to constitutive GFP expression (see Figure, **b**). After tamoxifen treatment, ~80% of the cardiomyocytes in the transgenic animal became GFP⁺, and 20% remained β -gal⁺GFP⁻. As in the analogous zebrafish experiment, any progenitor cells should remain GFP⁻ after the tamoxifen pulse. During normal ageing for up to 1 year after tamoxifen treatment, the ratio of GFP⁺ to β -gal⁺ cardiomyocytes remained fixed at 80:20, indicating no significant cardiomyocyte renewal by unlabelled progenitor cells. However, after infarction, the ratio of GFP⁺ to β -gal⁺ shifted to ~65:35 in the peri-infarct zone,

indicating that newly differentiated cardiomyocytes (β -gal⁺GFP⁻ because they had not undergone Cre-mediated recombination) had been recruited from the progenitor pool. Hence, in mice, the small amount of regeneration that occurs after injury involves the cardiac induction of progenitor cells.

a Zebrafish**b Mouse**

Limited regeneration in rodent hearts

Although they lack the remarkable regenerative capacity of the zebrafish heart, postnatal mammalian hearts also undergo some degree of cardiomyocyte renewal during normal ageing and disease. Despite all the recent attention by the field, this is not a new concept. Extremely low but detectable levels of cardiomyocyte cell-cycle activity have been reported in rodent studies dating back to the 1960s^{3,14}. Capturing the rare dividing cardiomyocytes present in mammalian hearts is technically challenging, but recent work has taken advantage of the greater specificity and throughput afforded by transgenic mouse models. For example, transgenic mice were created in which the cardiomyocyte-specific α -myosin heavy chain (*Myh6*, also known as α -MHC) promoter drives nuclear-localized expression of β -galactosidase¹⁵. This convenient read-out allowed researchers to screen more than 10,000 cardiomyocyte nuclei in histological sections for the incorporation of radiolabelled thymidine, and they found labelling indices of 0.0006%

for adult ventricular cardiomyocytes in intact hearts and 0.0083% for cardiomyocytes in the border zone of injured hearts^{14,15}.

Although such proliferative indices are small, they raise the possibility that such phenomena could be augmented therapeutically. Proof of concept for this approach has come from transgenic mice with cardiomyocyte-restricted overexpression of the cell-cycle activator cyclin D2, because these animals show reduced scar tissue and improved mechanical function after myocardial infarction¹⁶. Other efforts to enhance the proliferation of adult cardiomyocytes — by manipulating oncogenes or cell-cycle regulators — have proven less consistent in improving outcomes after infarction (see ref. 17 for a comprehensive review). Pharmacological enhancement of cardiomyocyte cell-cycle activity would be more practical clinically than gene therapy, and the signalling molecules periostin¹⁸, fibroblast growth factor-1 (ref. 19) and neuregulin 1 (NRG1)²⁰ have all been reported to act as mitogens for adult ventricular cardiomyocytes and to exert beneficial effects on

cardiac structure and function after infarction. (It should be noted that a more recent study has called into question the effects of periotin on cell-cycle activity or cardiac repair²¹.) A recent study of the mitogenic effects of NRG1 showed that simple systemic injection of this growth factor into adult mice enhanced infarct scar shrinkage and improved mechanical function²⁰. The effects of NRG1 on the cell cycle were dependent on the expression of its tyrosine kinase receptor, ERBB4, by cardiomyocytes and seem to stimulate mononucleated, but not binucleated, cardiomyocytes to divide. Although this intriguing result awaits independent confirmation, it suggests a straightforward approach to enhancing ventricular repair through the administration of recombinant growth factors.

Most of the above studies focused on the proliferation of existing cardiomyocytes, and were not designed to detect cardiomyocytes formed from progenitor cells. To determine whether such progenitor cells contribute to cardiomyocyte renewal, researchers have performed an elegant genetic fate-mapping experiment in transgenic mice²², akin to those previously described in the zebrafish model, in which cardiomyocytes were indelibly labelled after a tamoxifen pulse (Box 1). This system allowed the authors to distinguish between cardiomyocyte renewal from pre-existing (and therefore fluorescently labelled) cardiomyocytes and cardiomyocyte renewal from unlabelled progenitor cells. Interestingly, they found no significant contribution by such progenitor cells during normal ageing, up to one year after tamoxifen treatment. However, they observed a reduction in the fraction of labelled cardiomyocytes after infarction, indicating dilution by unlabelled progenitor cells. When combined with the findings that the rate of cardiomyocyte proliferation is very low in both normal and injured rodent hearts, these data indicate that the limited endogenous reparative mechanisms in the adult mammalian heart operate differently from those in zebrafish, and depend more on replenishment by cardiomyogenic progenitor cells than on replacement by cardiomyocyte proliferation.

A recent report suggests that these differences between mammalian and fish hearts do not necessarily apply earlier in development²³. Borrowing approaches from the zebrafish model, the authors resected the left ventricular apex of one-day-old neonatal mice and observed a brisk regenerative response reminiscent of that in the adult zebrafish. By three weeks after injury, the defect had been replaced by normal myocardial tissue, which showed normal contractile function by eight weeks. Genetic fate-mapping studies indicated that this regeneration was mediated by the proliferation of pre-existing cardiomyocytes, again as in the zebrafish. Notably, this regenerative capacity was not observed in seven-day-old mice, suggesting that its loss may coincide with cardiomyocyte binucleation and reduced cell-cycle activity. Nonetheless, in addition to representing a surgical tour de force, this study indicates that zebrafish-like regenerative mechanisms are latent in mammalian hearts. It also provides a genetically tractable model for dissecting the blocks to these mechanisms in the mammalian adult.

The evidence for human heart regeneration

Before addressing whether new cardiomyocytes are generated in the human heart after injury, it is instructive to review a few points about normal cardiac growth and adaptation to workloads (Box 2). In brief, most human cardiomyocyte nuclei are polyploid by the onset of puberty²⁴. In response to pathological workloads, such as hypertension, valvular disease and post-infarction overload, human cardiomyocytes commonly reinitiate DNA synthesis without nuclear division^{24,25}. This increases cardiomyocyte nuclear ploidy further, reaching levels as high as $64n$ (in which n represents the haploid set). Unlike rodent cardiomyocytes²⁶, most human cardiomyocytes seem to remain mononucleated throughout life²⁷. Thus, DNA synthesis is common in the adult human heart, but this cannot be equated to cardiomyocyte proliferation without accounting for the process of polyploidization.

Historically, regenerative responses have been detected by either the macroscopic regrowth of the tissue or the microscopic presence

of mitosis. Macroscopic regeneration of the human heart clearly does not occur. Mitosis occupies only ~2% of the cell cycle, making it hard to quantify meaningfully. Experiments have confirmed this, with some investigators reporting no mitosis after injury, and others reporting rare (and potentially abnormal) mitotic figures around the injured site^{28–30}. One factor contributing to these discrepancies has been the inherent difficulty in recognizing cardiomyocyte nuclei in conventional histological sections¹⁴. Many of the published images are persuasive for cardiomyocyte mitosis and provide important evidence that this can occur in humans. However, extrapolation to organ turnover rates from such low numbers is perilous.

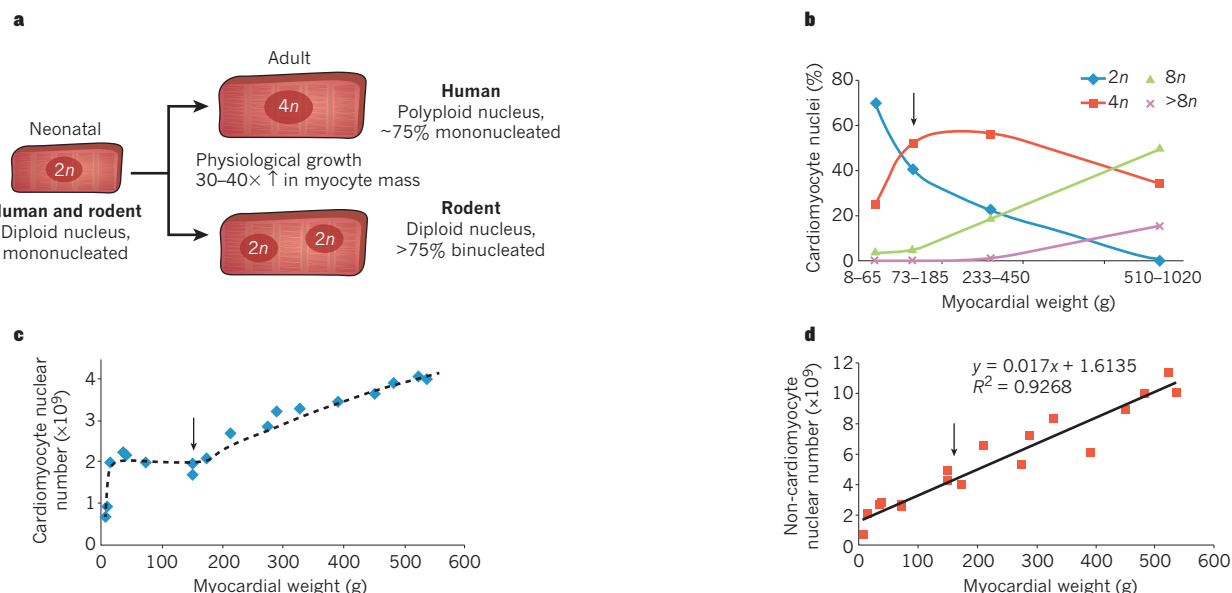
Several investigators have taken the approach of counting the number of cardiomyocytes in the human heart during normal and pathological growth^{24,25,31,32}, but this method is surprisingly difficult and requires many assumptions. Using a combination of meticulous dissection, histopathology, biochemical measurements of tissue DNA content and fluorescent analysis of individual nuclear DNA content, researchers have shown that from myocardial weights of 50–350 g, the cardiomyocyte nuclear number is steady at ~2 billion. Beyond that, there is a linear increase in nuclear number with increasing heart weight, reaching 4 billion cardiomyocyte nuclei in hypertrophied hearts weighing 700–900 g. The number of non-cardiomyocytes such as fibroblasts and vascular cells increases linearly with heart weight throughout life. If correct, these data indicate that cardiomyocyte renewal occurs during pathological hypertrophy. An important caveat is that the assignment of cardiomyocyte versus non-cardiomyocyte nuclear identity was based on the size and morphology of isolated nuclei. Because we know that postnatal growth and pathological hypertrophy are accompanied by increases in nuclear ploidy (and hence in size), it is possible that diploid cardiomyocyte nuclei in smaller hearts were mistakenly classified as non-cardiomyocyte nuclei. Increases in cardiomyocyte nuclear number in pathological hypertrophy have also been reported³³, using histological sections in which the cardiomyocyte nuclei can be more readily identified.

Two studies have attempted to use more direct means to measure the rates of cardiomyocyte DNA synthesis in human hearts. The first approach, by Bergmann *et al.*³⁴, was based on the worldwide pulse of ¹⁴C that occurred during the atmospheric testing of nuclear weapons in the cold war. The atmospheric ¹⁴C became incorporated into plants and entered the human food chain, labelling the DNA of dividing cells. After the Limited Nuclear Test Ban Treaty of 1963, atmospheric ¹⁴C levels dropped rapidly. This provided researchers with pulse-chase conditions that can be used to date cells, simply by identifying when atmospheric ¹⁴C levels match those of the DNA. As expected, non-cardiomyocytes in the normal human heart were found to be substantially younger than the patient, with ~18% turnover per year and a mean age of only four years. Notably, DNA from isolated cardiomyocyte nuclei (sorted by nuclear troponin staining) was also younger than the patient, although not nearly as young as that from non-cardiomyocytes. As indicated earlier, before one can infer cell division, it is essential to rule out a contribution from polyploidization. To do this, the authors sorted cardiomyocyte nuclei by DNA content, and analysed only the diploid DNA subset. The diploid cardiomyocyte nuclei were also younger than the patient, providing good evidence for cardiomyocyte division. Mathematical modelling suggested that cardiomyocyte renewal was age-dependent, with ~1% of cardiomyocytes being renewed per year at age 20, and 0.4% at age 75. On the basis of these kinetics, ~45% of cardiomyocytes would be predicted to be renewed over a normal human lifespan, whereas 55% would be cells persisting since birth.

In the second approach, by Kajstura *et al.*³⁵, the rates of cardiac DNA synthesis were obtained by examining post-mortem hearts from patients with cancer that had been treated with the thymidine analogue iododeoxyuridine (IdU). This agent is incorporated into nascent DNA, where it sensitizes cells to radiation therapy. IdU was given as bolus injections or multiweek infusions, and the time between treatment and death ranged from 7 days to 4.3 years. Using immunohistochemistry to

BOX 2

Nuclear dynamics during human heart growth



Fetal hearts in humans and rodents grow through the proliferation of mononucleated cardiomyocytes with diploid nuclei. In the first few days after birth, rodent cardiomyocytes withdraw from the cell cycle. By contrast, human cardiomyocytes seem to proliferate for the first few months after birth, after which replication slows markedly. The cells of both rodents and humans then undergo a period of physiological growth, increasing in size by 30–40-fold. Although most cardiomyocytes cannot grow this much with a single diploid genome, different species have taken varying approaches to solve this problem. Nearly all rodent cardiomyocytes undergo a final round of DNA replication followed by nuclear division without cytokinesis, resulting in a heart with more than 75% binucleated cells with a normal diploid (2n) content of DNA in each nucleus (see Figure, **a**)²⁶. In humans and other primates, most cardiomyocytes undergo a final round of DNA replication, without nuclear division or cytokinesis, resulting in mononucleated cells²⁷ with tetraploid (4n) or higher DNA content (**a**)³¹. (Estimates of human cardiomyocyte binucleation rates range from 25% in enzymatically dispersed fresh tissue²⁷ to more than 60% in potassium-hydroxide-digested formaldehyde-fixed tissue⁹⁹. We favour the 25% value, because the harsh potassium hydroxide digestion may selectively eliminate smaller mononucleated cardiomyocytes.) Notably, pacemaker cells of the sinoatrial and atrioventricular nodes remain small and diploid throughout life¹⁰⁰.

As the myocardial mass of the human heart increases, the percentage of diploid cardiomyocyte nuclei decreases steadily. This has been

demonstrated by cytofluorometric analysis of human cardiomyocyte nuclear DNA content as a function of myocardial weight, after carefully removing valves, vessels and fat (see Figure, **b**). (Hearts were grouped by weight into four bins for clarity.) Studies of paediatric human hearts indicate that polyploidization occurs in the pre-adolescent growth phase, from 8 to 12 years of age. Tetraploid nuclei (4n) are most common in the adult heart. During cardiac hypertrophy, octaploid nuclei (8n) become most common, with a substantial number of 16n nuclei or those with higher polyploidy. These data demonstrate that human cardiomyocytes have a substantial capacity for DNA replication.

Morphometric analysis has shown that the normal human adult number of 2 billion cardiomyocyte nuclei is reached by about 2 months of age. During physiological hypertrophy, the cardiomyocyte nuclear number remains steady. However, when the heart weight exceeds approximately 450 g (myocardial weight roughly 210 g), there seems to be a linear increase in cardiomyocyte nuclear number with increasing cardiac mass (see Figure, **c**). Because human hearts do not change nuclear number with hypertrophy, this is evidence for the generation of new cardiomyocytes, either from pre-existing cardiomyocytes or from stem cells. Non-cardiomyocyte nuclei increase linearly with increasing myocardial mass (see Figure, **d**), indicating that proliferation of these cells accompanies all phases of cardiac growth. The arrows in **b–d** denote the upper limits of normal for human myocardial weight. The data in panels **b–d** are derived from ref. 31; the trend line in **c** is hand drawn for illustration purposes only.

detect the IdU signal and identify cardiomyocytes, the researchers found remarkably high rates of cardiomyocyte DNA labelling, ranging from 2.5% to 46%. No IdU staining was found in control hearts from patients without cancer who had not been exposed to the radiosensitizer. Mathematical modelling suggested that cardiomyocytes turn over at a rate of 22% per year, compared with 20% for fibroblasts and 13% for endothelial cells. Furthermore, 83% of the cardiomyocyte nuclei were reported to be diploid, suggesting that this turnover reflects cell division, not increased nuclear ploidy.

It is hard to reconcile these two studies, which differ by nearly 50-fold in their estimates of cardiomyocyte turnover. An important difference seems to be related to the higher rates of cardiomyocyte DNA synthesis

activity in IdU-treated patients with cancer. Kajstura *et al.*³⁵ reported threefold lower DNA synthesis rates (based on immunolabelling of the cell proliferation marker Ki-67) in control hearts from patients without cancer than in hearts from IdU-treated patients with cancer. Neither study adequately rules out a contribution from DNA repair, which can masquerade as DNA replication in these assays. This is of particular concern in patients with cancer receiving radiation treatment plus a radiosensitizer. Kajstura *et al.* suggested that only senescent cardiomyocyte nuclei contain troponin, which could bias the turnover studies of Bergmann and his colleagues³⁴ towards low proliferation. However, a follow-up paper by Bergmann *et al.*³⁶ provided evidence that nearly all cardiomyocyte nuclei were identified by troponin

staining. Furthermore, the findings of Kajstura *et al.* contradict two well-accepted principles. First, the findings suggest that more than 80% of cardiomyocyte nuclei are diploid, in contrast to most other reports that suggest they are polyploid. If the nuclei were in fact polyploid, then polyploidization could underlie the authors' high estimates of DNA synthesis. Second, the authors conclude that cardiomyocytes are as proliferative as non-cardiomyocytes, whereas most other investigators find greater orders of magnitude of proliferation in non-cardiomyocytes. Indeed, the cardiomyocyte IdU-incorporation rate (2.5–46%) detected by Kajstura *et al.*³⁵ approaches the rate reported previously for sarcomas targeted by IdU (50–70%)³⁷. The heart does not proliferate like a sarcoma, so these cardiac IdU-incorporation estimates must be too high.

Taken together, these human studies provide strong evidence for plasticity in the adult human heart. There is extensive morphometric evidence for DNA synthesis and an increase in cardiomyocyte number in diseased human hearts. Cardiomyocyte division or generation from progenitor cells probably occurs in the human heart, but it seems to be a very slow process. We need better tools to study this process quantitatively, and better ways to model it, if we hope to exploit it therapeutically.

Stem cells and cell therapy

Stem-cell biology is one of the fastest moving areas of biomedical research, and among all of the solid organs, the heart has one of the most active regeneration research programmes. The field can be conceptually organized into work involving endogenous and exogenous cells. The many exogenous cell types can be further divided into pluripotent cells (such as embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs)) and adult cells of more limited potential (such as circulating progenitor cells, resident cardiac progenitor cells and cells native to other tissues). Here we focus on the cells closest to clinical trials and those for which there are the most reliable data.

Cardiac progenitor cells

Several investigators have reported resident populations of cardiac progenitor cells (CPCs) in postnatal hearts. These were identified using a variety of approaches, including studying the expression of surface markers such as c-KIT or SCA-1 (also known as LY6A; note that SCA-1 has no apparent human orthologue) and physiological properties such as the ability to efflux fluorescent dye or form multicellular spheroids (reviewed in refs 38 and 39). Initially, it seemed that there was little overlap among CPCs identified by the different methods, and some scientists suggested that several populations of CPC exist. More recent studies indicate shared markers among once-distinct populations or different stages of maturation in the same line of cells^{40,41}, so the field may be converging.

CPCs expressing the tyrosine kinase receptor c-KIT are the most extensively studied. In the human adult, c-KIT is expressed by telocytes (formerly known as the interstitial cells of Cajal), the thymic epithelium and mature circulating cells such as haematopoietic stem cells and mast cells. Immature endothelial cells and cardiomyocytes also express c-KIT during development⁴². Small round cells expressing c-KIT have been identified in the perivascular compartment of the adult heart, and their abundance increases in human heart failure⁴³. After isolation from rat and human hearts, c-KIT⁺ cells have been reported to give rise to cardiomyocytes, smooth muscle cells and endothelial cells. Some studies indicate that, when transplanted, c-KIT⁺ cells induce large-scale regeneration of myocardial infarcts and contribute to the formation of new myocardium and vessels⁴⁴, whereas others suggest smaller-scale regeneration⁴⁵. On the basis of these data, a clinical trial is under way, testing the safety and feasibility of autologous c-KIT⁺ cells as an adjunctive treatment for patients undergoing coronary bypass surgery (ClinicalTrials.gov identifier NCT00474461).

Not all studies with c-KIT⁺ CPCs gave robustly positive results. In studies with genetic read-outs for lineage tracing and differentiation state, c-KIT⁺ cells from the adult mouse heart have not been shown

to differentiate into cardiomyocytes *in vitro* or after transplantation into infarcted hearts⁴⁶. Another study using transgenic reporter mice found no evidence to suggest that endogenous c-KIT⁺ cells differentiate into cardiomyocytes, although re-expression of c-KIT in pre-existing cardiomyocytes was identified after injury⁴². Others point out that myocardium, like all solid tissues, contains mast cells. Mast cells are small round cells that reside in clusters in the perivascular space, strongly express c-KIT and increase in number in failing hearts⁴⁷. Studies in humans suggest that 90–100% of all of the cardiac c-KIT⁺ cells are actually mast cells. However, expansion in culture seems to select for c-KIT⁺ cells that lack mast-cell markers, indicating that freshly isolated cells and cultured cells are different populations⁴⁸.

Another CPC population in clinical trials is cardiosphere-forming cells. These cells are isolated on the basis of their ability to migrate out of cultured cardiac tissue fragments and form spheroids in suspension cultures^{49,50}. As one might predict, this yields a mixture of cells, some of which express stem-cell markers such as c-KIT, and others that seem to come from the stromal-vascular compartment. CPCs have been reported to give rise to cardiomyocytes *in vitro* and *in vivo* after transplantation, and to enhance cardiac function after infarction⁵⁰. On the basis of these data, a clinical trial of autologous CPCs has been initiated for patients with recent myocardial infarctions (NCT00893360). The 'stemness' of CPCs has recently been questioned, and it has been suggested that these cells are principally cardiac fibroblasts and that CPC-derived cardiomyocytes are contaminants derived from the original tissue⁵¹.

Thus, although the study of CPCs is an exciting, new area of cardiac research, it is also one of the most controversial. Most of the work has focused on cell culture and transplantation, driven by the clinical need for cardiac repair. We know almost nothing about the endogenous behaviour of CPCs, however. An important question remains about the role of these cells in development, homeostasis, ageing and reaction to injury. The field needs models that permit unambiguous tracing of CPC lineage and phenotype without resorting to transplantation or cell culture (Box 1).

Bone marrow cells

Considerable interest in bone-marrow-derived cells for cardiac repair was prompted by reports of haematopoietic stem cells transdifferentiating into cardiomyocytes⁵². Subsequent studies have shown that haematopoietic stem cells do not form cardiomyocytes but instead become mature blood cells after transplantation^{53,54}. Nevertheless, animal studies show improvements in ventricular function when haematopoietic cells are administered after infarction, implicating paracrine signalling as the major mechanism of action.

Work with marrow-derived stromal cells (MSCs) has followed a similar trajectory. MSCs were originally reported to transdifferentiate into cardiomyocytes⁵⁵ but are now thought to exert their main actions in a paracrine manner through the release of cytokines⁵⁶. Interestingly, most MSCs die within days or weeks of transplantation into infarcts, yet their beneficial effects can be seen long term, suggesting a critical window of time for the action of MSCs after infarction. MSCs probably operate by many mechanisms, but considerable evidence points towards regulation of the WNT pathway. MSCs secrete antagonists of canonical WNT ligands, such as secreted frizzled related protein 2 (ref. 56). Blocking the production of WNT antagonists limits the beneficial effects of mouse MSCs. A recent report has shown that the administration of MSCs to pig infarcts stimulated endogenous CPCs to contribute to the repair of the infarcts⁵⁷. Further identification of paracrine mediators may allow the development of simpler, cell-free treatments based on proteins or small molecules.

Clinical trials have mostly focused on the delivery of bone marrow mononuclear cells by the coronary circulation. It should be emphasized that >99.9% of bone marrow mononuclear cells are not stem cells, but are committed, although immature, granulocytes or other haematopoietic lineages. These trials indicate that the delivery of bone marrow

derivatives through the coronaries is feasible and safe, but the benefits are modest. MSCs are also in clinical trials (NCT00587990). There are few published results with these cells, but one of the strongest cardiac-repair treatment effects seen so far (a 14% improvement in ejection fraction — the fraction of blood ejected from the left ventricle during one contraction) was reported after the intracoronary administration of large numbers of autologous MSCs⁵⁸. Allogeneic MSCs administered to patients intravenously within ten days of infarction were well tolerated and were associated with decreased arrhythmias and an improvement in some indices of contractile function⁵⁹.

Taken together, the best current evidence indicates that bone marrow cells do not work by directly differentiating into new cardiomyocytes. Instead, the cells have been shown to elaborate signals that control the response of cells native to the myocardium, and thereby regulate healing. Although many view this as a novel aspect of stem-cell biology, students of pathology will recognize that this phenomenon fits under a more familiar heading: inflammation. We find it useful to consider the participation of marrow derivatives in cardiac repair as part of the inflammatory response, which is known to regulate angiogenesis, cardiomyocyte survival and left ventricular remodelling after infarction.

Pluripotent stem cells

Many types of adult stem cell are unable to generate large numbers of unambiguous cardiomyocytes. This limitation does not apply to ESCs or their more recently developed 'man-made' counterpart, iPSCs. Because both ESCs and iPSCs can be propagated indefinitely, while still retaining the capacity to differentiate into almost all cell types, they are a potentially inexhaustible supply of human cardiomyocytes. Our current thinking about how cardiomyocytes arise from ESCs is shown in Fig. 1. Human ESC-derived cardiomyocytes express early cardiac transcription factors such as NKX2.5, as well as the expected sarcomeric proteins, ion channels, connexins and calcium-handling proteins (Fig. 2). They show similar functional properties to those reported for cardiomyocytes in the developing heart, and undergo comparable

mechanisms of excitation–contraction coupling and neurohormonal signalling^{60–63}. Although human ESC-derived cardiomyocytes have been more intensively studied, data indicate that human iPSC-derived cardiomyocytes have a very similar phenotype^{64,65}. Importantly, cardiomyocytes from either pluripotent stem-cell type are immature and so lack the expression profile, morphology and function of adult ventricular cardiomyocytes.

The cardiac potential of ESCs and iPSCs is indisputable, but their unique origin and pluripotency presents a new set of challenges. ESCs are derived from the inner cell mass of preimplantation-stage blastocysts⁶⁶, and this contributes to the ethical controversy surrounding their use. Moreover, ESC-based therapies will be allogeneic and require immunosuppression. iPSCs were originally generated by the reprogramming of adult somatic cells such as dermal fibroblasts by the forced expression of up to four stem-cell-related transcription factors^{67–69}. As such, their derivation does not involve the destruction of embryos, and they could be used in autologous cell therapies. Nonetheless, first-generation iPSCs were problematic because the reprogramming factors were introduced using integrating viruses, raising concerns about neoplastic transformation. More recently, there have been a variety of refinements to iPSC generation that should reduce or eliminate this risk, including the use of episomal gene delivery, excisable transgenes, cell-permeable recombinant proteins and synthetic messenger RNA (see ref. 70 and references therein). Perhaps most notably, several small molecules have been shown to greatly enhance the efficiency of reprogramming⁷¹, inviting speculation that iPSCs may be generated using such factors alone in the near future. Further work will be required to more precisely define the phenotype and maturation potential of cardiomyocytes derived from iPSCs generated by these methods.

Another concern relating to the clinical application of pluripotent stem cells is their capacity to form teratomas after transplantation⁷². To overcome this, the field needs to develop methods to enrich ESC and iPSC derivatives for cardiomyocytes or other useful cell types

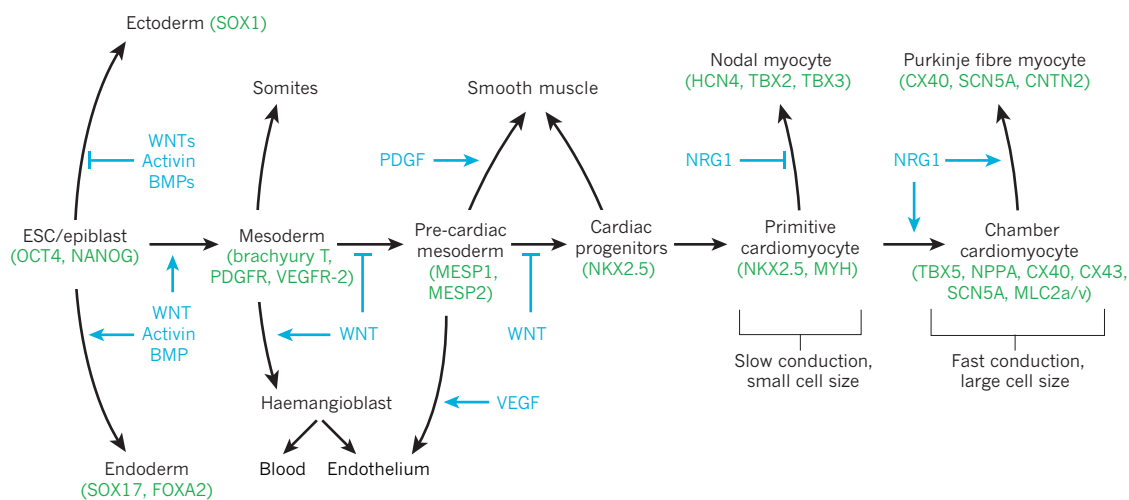


Figure 1 | Cardiovascular lineages during embryonic development and ESC differentiation. Cardiac differentiation from ESCs closely mimics cardiac development in the embryo. In either case, the specification of the cardiovascular lineages involves a transition through a sequence of increasingly restricted progenitor cells, proceeding from a pluripotent state to mesoderm and then to cells committed to cardiovascular fates. Growth factors that regulate fate choices are listed at branch points (blue), and key transcription factors and surface markers for each cell state are listed under the cell types (green). The growth factors are useful for directing the differentiation of ESCs, whereas the markers are useful for purifying cells at defined developmental states. Primitive cardiomyocytes in the embryonic heart tube and nodal or pacemaker cells show slow electrical propagation and a small cell size. By contrast, the eventual specification of working atrial and ventricular cardiomyocytes is accompanied

by more rapid conduction, ion-channel remodelling and increased cell size. Although the field has made considerable progress towards determining the early events of cardiogenesis, a better understanding of how pacemaker and chamber-specific cardiac subtypes are formed is required for clinical applications. BMPs, bone morphogenetic proteins; CNTN2, contactin-2; CX, connexin; FOXA2, forkhead box protein A2; HCN4, potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 4; MESP, mesoderm posterior protein; MLC2a/v, myosin light chain 2a and/or 2v; MYH, myosin heavy chain; NPPA, natriuretic peptide precursor A; NRG1, neuregulin 1; PDGF, platelet-derived growth factor; PDGFR, PDGF receptor; SCN5A, sodium channel protein type 5 subunit α ; SOX, SRY-related high-mobility-group box; TBX, T-box transcription factor; VEGF, vascular endothelial growth factor; VEGFR-2, VEGF receptor-2.

(such as endothelial, smooth muscle and stromal cells). Historically, ESCs have been differentiated by culture in three-dimensional aggregates known as embryoid bodies, in medium containing a high percentage of fetal calf serum. This method is poorly cardiogenic, and differentiated human embryoid bodies are typically composed of less than 1% cardiomyocytes⁷³. More recently, our group and others have used insights from developmental biology to devise better controlled approaches in which human ESCs and iPSCs are treated with defined factors, resulting in highly enriched populations of cardiomyocytes^{74–76}. A common theme with such methods has been the manipulation of cardioinductive molecules belonging to the transforming growth factor- β superfamily — specifically, activin and the bone morphogenetic proteins (BMPs). Our group has reported a protocol involving the serial application of activin A and BMP4, for example, which reliably yields ~60–80% human ESC-derived cardiomyocytes in large-scale preparations (~ 10^8 – 10^9 total cells)^{74,77}. Further refinements are possible by manipulating the WNT- β -catenin signalling pathway⁷⁸, which mediates biphasic effects on ESC cardiogenesis, promoting mesodermal induction early but inhibiting cardiogenesis late⁷⁹.

A complementary approach involves the isolation of mesodermal progenitor cells with a more restricted potential, such as cardiovascular progenitor cells that can differentiate into cardiomyocytes, smooth muscle cells and endothelial cells. Such multipotent progenitor cells have been identified in differentiating ESC cultures on the basis of their expression of transcription factors such as mesoderm posterior protein 1 (MESP1)⁸⁰, NKX2.5 (ref. 81) and ISL1 (refs 82 and 83). Arguably more useful for eventual clinical application are progenitor populations that can be sorted on the basis of their expression of a cell-surface marker, such as the cardiovascular progenitor cells marked by expression of vascular endothelial growth factor receptor-2 (VEGFR-2, also known as FLK1 and KDR)⁷⁵. If such cells could be induced to self-renew, they would potentially be very useful for cardiac repair.

Human ESC-derived cardiomyocytes have been shown to engraft in infarcted mouse, rat, guinea pig and pig hearts (Fig. 3), forming islands of nascent, proliferating human myocardium within the scar zone^{74,84,85}. This partial remuscularization was accompanied by beneficial effects on regional and global cardiac function^{74,84}, although some investigators have questioned whether these effects are sustained at later time points⁸⁶. Notably, the mechanism (or mechanisms) underlying the observed improvements in contractile function remains unresolved. In the aforementioned rodent studies, most of the graft tissue was isolated from the host myocardium by means of scar tissue, which may prevent synchronous beating. Furthermore, these human cells, which fire *in vitro* at ~50–150 beats per minute (b.p.m.)⁷⁷, may not keep pace with the rapid rate of rats (~400 b.p.m.) and mice (~600 b.p.m.). If they cannot, then the observed salutary effects probably resulted from an indirect, paracrine mechanism, like those described above for adult cells. This also indicates that further beneficial effects on cardiac function may be possible after transplantation to a slower-rated recipient, such as a canine or porcine infarct model.

Reprogramming fibroblasts to cardiomyocytes

Fifteen years ago, researchers showed that fibroblasts could be transdifferentiated into skeletal muscle *in vitro* or in the injured heart by overexpressing the gene encoding the myogenic transcription factor, MyoD⁸⁷. Despite an intensive search by several groups, no comparable master gene for cardiac muscle was found, and interest in reprogramming waned. Spurred by the discovery of iPSCs, scientists have returned to this field, using combinations of transcription factors to reactivate core transcriptional networks of desired cell types. In an attempt to induce cardiac differentiation, researchers performed a systematic screen of 14 cardiac transcription factors for their ability to activate a cardiac-specific transgene — the *Myh6* promoter driving yellow fluorescent protein (YFP) expression — in cardiac fibroblasts⁸⁸. The full cocktail activated fluorescence in ~1% of cells. A systematic

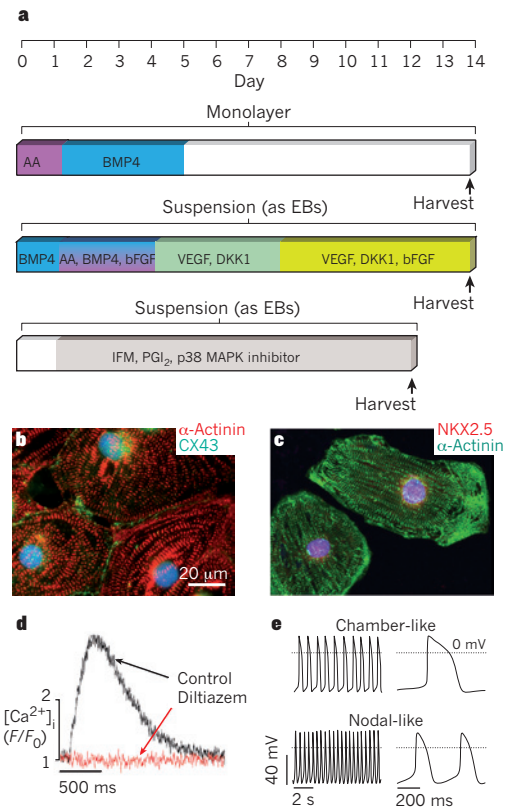


Figure 2 | Guided differentiation and phenotype of cardiomyocytes from pluripotent stem cells. **a**, Selected protocols for the guided differentiation of human ESCs and iPSCs into cardiomyocytes using chemically defined factors. The top timeline shows a protocol from our group in which differentiating cells are serially pulsed with activin A (AA) and BMP4 under monolayer culture conditions⁷⁴. The middle timeline shows a protocol from ref. 75 that involves embryoid body (EB) formation in suspension cultures, and the application of several signalling molecules such as activin A, BMP4, basic fibroblast growth factor (bFGF), dickkopf-related protein 1 (DKK1) and VEGF. The bottom timeline shows a protocol from ref. 76, in which embryoid bodies in suspension are continuously cultured in insulin-free medium (IFM) supplemented with prostaglandin I₂ (PGI₂) and an inhibitor of p38 MAP kinase (MAPK). **b**, Representative human ESC-derived cardiomyocytes, differentiated using the monolayer protocol (top timeline in **a**), immunostained for α -actinin (red) and CX43 (green). Nuclei are shown in blue. **c**, Representative human iPSC-derived cardiomyocytes, differentiated using the monolayer protocol (top timeline in **a**), immunostained for α -actinin (green) and the transcription factor NKX2.5 (red). **d**, Intracellular $[Ca^{2+}]_i$ transients in a human ESC-derived cardiomyocyte before (black) or after (red) the application of diltiazem, an L-type Ca^{2+} -channel blocker. The absence of $[Ca^{2+}]_i$ transients after diltiazem treatment indicates that extracellular Ca^{2+} is required to initiate intracellular Ca^{2+} release, just as in adult cardiomyocytes. F/F_0 denotes the change in fluorescence intensity. **e**, Human ESC-derived cardiomyocytes show the characteristic action-potential properties of either working chamber (top) or nodal (bottom) cardiomyocytes, indicating early subtype specification.

winnowing yielded three transcription factors (MEF2C, GATA4 and TBX5) that activated the transgene in 20% of fibroblasts. About 4% of the cells expressed endogenous sarcomeric proteins such as cardiac troponin T, and only ~1% showed functional properties such as spontaneous beating. Thus, most of the YFP⁺ cells were only partially reprogrammed, although their global gene expression patterns had shifted markedly from fibroblast to cardiomyocyte.

While this manuscript was under review, a different method of reprogramming mouse embryonic fibroblasts to cardiomyocytes was reported⁸⁹. This group used the ‘Yamanaka factors’ — OCT4 (also known as POU5F1), SOX2, KLF4 and c-MYC — to initiate

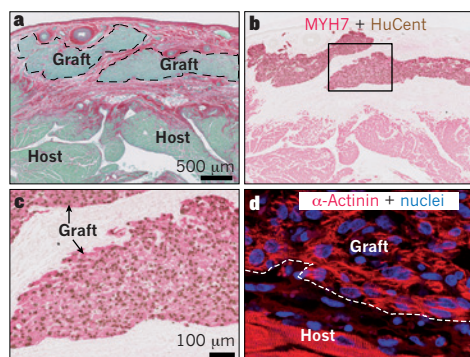


Figure 3 | Grafts of human ESC-derived cardiomyocytes in the cryoinjured guinea-pig heart. Representative photomicrographs demonstrating substantial implants of human myocardium within the scar tissue. **a**, Using picrosirius red stain, the scar appears red, and viable tissue is green. **b**, The human origin of the graft myocardium was confirmed in an adjacent section by combined *in situ* hybridization, with a human-specific pan-centromeric (HuCent; brown) probe, and β -myosin heavy chain (MYH7; red) immunohistochemistry. **c**, Inset from **b** at higher magnification. The nuclear localization of the HuCent signal confirms the human origin of these cells. **d**, Immunostaining for α -actinin (red) highlights the sarcomeric organization of the graft cardiomyocytes. Nuclei have been counterstained with Hoechst 33342 (blue).

reprogramming, but they blocked signalling through the JAK–STAT pathway, which is required for pluripotency in the mouse, and added the cardiogenic factor BMP4. These modifications yielded minimal generation of iPSCs, but instead activated the cardiac progenitor program and, within 2 weeks, generated substantial numbers of beating colonies. By 18 days after induction, approximately 40% of the cells expressed cardiac troponin T. The authors attributed the increased efficiency to the generation of highly proliferative progenitor cells, as opposed to the formation of cardiomyocytes with low proliferative potential. It should also be noted that this study used mouse embryonic fibroblasts, whereas the systematic screen of 14 transcription factors was principally in postnatal mouse cardiac fibroblasts.

Reprogramming the scar-forming fibroblast to a cardiomyocyte is intuitively appealing, particularly if it can be done directly in the infarct. To succeed clinically, we need to know how normal these reprogrammed cardiomyocytes are, and the process will have to be much more efficient and transgene-free. Despite some challenges, this is an exciting avenue of research and could be a game changer.

Tissue engineering

Tissue engineering refers to the growth of three-dimensional tissues *in vitro*, with the aim of building more biologically relevant models for *in vitro* study or tissues for *in vivo* regenerative therapy. Most commonly, this involves the use of porous, biodegradable scaffolds onto which cells are seeded, but other approaches include casting cells into hydrogels or creating scaffold-free tissues composed only of cells and the matrix they secrete. Synthetic materials have big advantages in manufacturability, typically being easy, cheap and reproducible to make. However, synthetics generally have worse biocompatibility, because they cause foreign-body inflammatory reactions and, sometimes, release locally toxic degradation products. Bioreactors are often used in tissue engineering to provide electrical and mechanical conditioning or to deliver nutrients to the tissue by perfusion systems.

A major aim of bioengineering is to improve the host response to biomaterials, in essence, to make materials that can heal⁹⁰. Surprisingly, the chemical composition of a material does not have a major influence on how the body responds to it. Whether materials are organic or metallic, hydrophobic or hydrophilic, or positively or negatively charged, they all cause similar foreign-body reactions. Instead, what the body seems to sense is the surface topography of a material⁹⁰. When surfaces are smooth,

there is intense inflammation and scarring, creating a fibrotic capsule around the implant. If a surface is given a more complex topography, for example, by creating pores or grooves, there is less inflammation, scarring diminishes and blood vessels grow into the implant. Systematic variation in the topology can ‘tune’ this host response. For example, our tissue-engineering group has developed scaffolds with two compartments: cylindrical channels to generate cables of cardiomyocytes, surrounded by a network of smaller interconnected pores for stromal and vascular ingrowth⁹¹. The pores are optimally sized to maximize vascularization within the implant and minimize fibrosis around it.

The cardiomyocytes used in tissue engineering have been immature cells derived from young animals or stem cells. To take on an adult workload, these cells will need to organize into the cable-like structure of myocardium and increase their size by more than 20-fold compared with the neonatal stage. There is a continuing debate in tissue engineering about whether this maturation should take place before or after transplantation. On the one hand, electrical⁹² and mechanical⁹³ stimulation *in vitro* enhance hypertrophy, alignment and electromechanical function of rat cardiomyocyte constructs. On the other hand, greater cell differentiation is associated with worse survival after transplantation⁹⁴, so there is probably a point of diminishing returns. This needs to be explored further experimentally.

One of the big lessons from tissue engineering has come from studies comparing cardiomyocyte-only with mixed-cell constructs. When cardiomyocyte-only constructs are transplanted, the tissue survives poorly. When vascular endothelial cells together with a stromal cell population are included, the endothelial cells form networks resembling a primitive vascular plexus, and the stromal cells form a provisional matrix that enhances mechanical integrity^{95,96}. After transplantation, the endothelial network organizes into a definitive vascular network that connects to the host circulation, bringing blood flow into the tissue several days sooner than would otherwise be seen. Indeed, our group and others have demonstrated improved survival of prevascularized human myocardial constructs incorporating vascular and stromal elements compared with constructs containing cardiomyocytes alone^{95,97}. This indicates that there is considerable synergy to including vessels and connective tissue elements when engineering tissue.

Tissue engineering has not been as extensively studied as cell transplantation in preclinical disease models, but initial studies are promising. A recent study⁹³ prepared constructs of engineered rat heart tissue from neonatal cardiomyocytes and conditioned them for several days using a cyclic stretch system. The constructs were sutured to the surface of rat hearts that had been infarcted two weeks previously, and were studied one month after implantation. Compared with infarcted hearts receiving non-contractile constructs, hearts receiving the engineered heart tissue had better contractile function, and interestingly, conduction velocities across the infarct were improved, probably because the grafts had electrically connected to the surrounding viable myocardium. Another group reported that patches generated from cardiosphere-derived CPCs can enhance heart function after infarction⁹⁸, and there are hints that tissue engineering also provides a larger graft size compared with cell transplantation.

Perspective

After more than a decade of furious activity, the science of stem cells seems to be catching up with its promise. Clinical-scale preparations of the main cardiac cell types can now be generated, and we are learning the rules for building myocardium and keeping it alive after transplantation. Clinical trials have established techniques for cell delivery, and protocols for establishing feasibility, safety and early-stage efficacy in humans are in place. The first patient trials have demonstrated safety with hints of efficacy. So far, so good.

That said, many short- and long-term challenges remain. In the near term, it will be important to derive the right subtype of cardiomyocyte, for example, ventricular cardiomyocytes that are free of pacemaker cells for repair of an infarct. The major challenge facing the field of

adult CPCs is to develop protocols with higher yields of definitive cardiomyocytes. Researchers studying pluripotent stem cells need to identify the optimal stage of differentiation and demonstrate that these cells can be used without tumorigenesis. The question of allogeneic versus autologous cells remains open. Although desirable, autologous cells will be more expensive, more variable, and the time needed to expand them precludes their use in any acute setting. Allogeneic cells will provide the only off-the-shelf product, but we need to learn how best to manage the immune response to prevent their rejection. All of these efforts will be advanced by improvements in integration of the graft, including control of vascularization (growth of both arterial conduits and microvasculature), inflammation and scarring.

Further ahead, *in situ* manipulation of cells in the heart may allow us to control their fates, thereby obviating transplantation. For example, it may be possible to control CPCs using small molecules or growth factors to enhance their regenerative abilities. Fibroblasts in infarcts could potentially be reprogrammed directly to cardiomyocytes. Given our increasing ability to control the fates of cells and tissues, the debate over whether the heart is intrinsically terminally differentiated seems anachronistic, for the heart does not exist apart from the person who knows how to manipulate it. It is more useful to ask what we can do to promote cardiac regeneration best, and then do it. ■

- Carvalho, A. B. & de Carvalho, A. C. Heart regeneration: past, present and future. *World J. Cardiol.* **2**, 107–111 (2010).
- Rumyantsev, P. P. in *Muscle Regeneration* (ed. Mauro, A.) 335–355 (Raven, 1979).
- Rumyantsev, P. P. *Growth and Hyperplasia of Cardiac Muscle Cells* (Harwood Academic, 1991).
- Murry, C. E., Reinecke, H. & Pabon, L. M. Regeneration gaps: observations on stem cells and cardiac repair. *J. Am. Coll. Cardiol.* **47**, 1777–1785 (2006).
- Whelan, R. S., Kaplinsky, V. & Kitsis, R. N. Cell death in the pathogenesis of heart disease: mechanisms and significance. *Annu. Rev. Physiol.* **72**, 19–44 (2010).
- Olivetti, G., Melissari, M., Capasso, J. M. & Anversa, P. Cardiomyopathy of the aging human heart. *Circ. Res.* **68**, 1560–1568 (1991).
- Laflamme, M. A., Zbinden, S., Epstein, S. E. & Murry, C. E. Cell-based therapy for myocardial ischemia and infarction: pathophysiological mechanisms. *Annu. Rev. Pathol.* **2**, 307–339 (2007).
- Povsic, T. J. & O'Connor, C. M. Cell therapy for heart failure: the need for a new therapeutic strategy. *Expert Rev. Cardiovasc. Ther.* **8**, 1107–1126 (2010).
- Oberpriller, J. O. & Oberpriller, J. C. Response of the adult newt ventricle to injury. *J. Exp. Zool.* **187**, 249–253 (1974).
- Poss, K. D., Wilson, L. G. & Keating, M. T. Heart regeneration in zebrafish. *Science* **298**, 2188–2190 (2002).
- Lepilina, A. *et al.* A dynamic epicardial injury response supports progenitor cell activity during zebrafish heart regeneration. *Cell* **127**, 607–619 (2006).
- Kikuchi, K. *et al.* Primary contribution to zebrafish heart regeneration by *gata4*⁺ cardiomyocytes. *Nature* **464**, 601–605 (2010).
- Jopling, C. *et al.* Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **464**, 606–609 (2010).
- References 12 and 13 demonstrate that heart regeneration in zebrafish is mediated by the proliferation of pre-existing cardiomyocytes, not by the mobilization of undifferentiated precursors.**
- Soonpaa, M. H. & Field, L. J. Survey of studies examining mammalian cardiomyocyte DNA synthesis. *Circ. Res.* **83**, 15–26 (1998).
- Soonpaa, M. H. & Field, L. J. Assessment of cardiomyocyte DNA synthesis in normal and injured adult mouse hearts. *Am. J. Physiol.* **272**, H220–H226 (1997).
- Hassink, R. J. *et al.* Cardiomyocyte cell cycle activation improves cardiac function after myocardial infarction. *Cardiovasc. Res.* **78**, 18–25 (2008).
- Ahuja, P., Sdek, P. & MacLellan, W. R. Cardiac myocyte cell cycle control in development, disease, and regeneration. *Physiol. Rev.* **87**, 521–544 (2007).
- Kuhn, B. *et al.* Perostin induces proliferation of differentiated cardiomyocytes and promotes cardiac repair. *Nature Med.* **13**, 962–969 (2007).
- Engel, F. B., Hsieh, P. C., Lee, R. T. & Keating, M. T. FGF1/p38 MAP kinase inhibitor therapy induces cardiomyocyte mitosis, reduces scarring, and rescues function after myocardial infarction. *Proc. Natl Acad. Sci. USA* **103**, 15546–15551 (2006).
- Bersell, K., Arab, S., Haring, B. & Kuhn, B. Neuregulin1/ErbB4 signaling induces cardiomyocyte proliferation and repair of heart injury. *Cell* **138**, 257–270 (2009).
- This study identifies neuregulin as a mitogen for cardiomyocytes, and shows that systemic NRG1 administration improves function in infarcted mouse hearts.**
- Lorts, A., Schwaneckamp, J. A., Elrod, J. W., Sargent, M. A. & Molkenstein, J. D. Genetic manipulation of perostin expression in the heart does not affect myocyte content, cell cycle activity, or cardiac repair. *Circ. Res.* **104**, e1–e7 (2009).
- Hsieh, P. C. *et al.* Evidence from a genetic fate-mapping study that stem cells refresh adult mammalian cardiomyocytes after injury. *Nature Med.* **13**, 970–974 (2007).
- These authors use transgenic mice and a genetic fate-mapping strategy to show that progenitor cells contribute to cardiomyocyte renewal after injury.**
- Porrello, E. R. *et al.* Transient regenerative potential of the neonatal mouse heart. *Science* **331**, 1078–1080 (2011).
- Adler, C. P. Relationship between deoxyribonucleic acid content and nucleoli in human heart muscle cells and estimation of cell number during cardiac growth and hyperfunction. *Recent Adv. Stud. Cardiac Struct. Metab.* **8**, 373–386 (1975).
- This study uses careful biochemical and cytophotometric techniques to determine cardiomyocyte number and DNA content in human hearts during normal ageing and disease.**
- Adler, C. P. & Friedburg, H. Myocardial DNA content, ploidy level and cell number in geriatric hearts: post-mortem examinations of human myocardium in old age. *J. Mol. Cell Cardiol.* **18**, 39–53 (1986).
- Li, F., Wang, X., Capasso, J. M. & Gerdes, A. M. Rapid transition of cardiac myocytes from hyperplasia to hypertrophy during postnatal development. *J. Mol. Cell Cardiol.* **28**, 1737–1746 (1996).
- Olivetti, G. *et al.* Aging, cardiac hypertrophy and ischemic cardiomyopathy do not affect the proportion of mononucleated and multinucleated myocytes in the human heart. *J. Mol. Cell Cardiol.* **28**, 1463–1477 (1996).
- MacMahon, H. E. Hyperplasia and regeneration of the myocardium in infants and in children. *Am. J. Pathol.* **13**, 845–854 (1937).
- Beltrami, A. P. *et al.* Evidence that human cardiac myocytes divide after myocardial infarction. *N. Engl. J. Med.* **344**, 1750–1757 (2001).
- Rumyantsev, P. P. in *Growth and Hyperplasia of Cardiac Muscle Cells* (ed. Carlson, B. M.) 210–238 (Harwood Academic Publishers, 1991).
- Adler, C. P. & Costabel, U. Cell number in human heart in atrophy, hypertrophy, and under the influence of cytostatics. *Recent Adv. Stud. Cardiac Struct. Metab.* **6**, 343–355 (1975).
- Herget, G. W., Neuburger, M., Plagwitz, R. & Adler, C. P. DNA content, ploidy level and number of nuclei in the human heart after myocardial infarction. *Cardiovasc. Res.* **36**, 45–51 (1997).
- Grajek, S. *et al.* Hypertrophy or hyperplasia in cardiac muscle. *Eur. Heart J.* **14**, 40–47 (1993).
- Bergmann, O. *et al.* Evidence for cardiomyocyte renewal in humans. *Science* **324**, 98–102 (2009).
- In this study, the authors took advantage of the global pulse of ¹⁴C during the cold war and estimated a rate of cardiomyocyte renewal in human hearts of up to 1% per year.**
- Kajstura, J. *et al.* Cardiomyogenesis in the adult human heart. *Circ. Res.* **107**, 305–315 (2010).
- Bergmann, O. *et al.* Identification of cardiomyocyte nuclei and assessment of ploidy for the analysis of cell turnover. *Exp. Cell Res.* **317**, 188–194 (2010).
- Kinsella, T. J. *et al.* A phase I study of intermittent intravenous bromodeoxyuridine (BUDR) with conventional fractionated irradiation. *Int. J. Radiat. Oncol. Biol. Phys.* **10**, 69–76 (1984).
- Segers, V. F. & Lee, R. T. Stem-cell therapy for cardiac disease. *Nature* **451**, 937–942 (2008).
- Passier, R., van Laake, L. W. & Mummery, C. L. Stem-cell-based therapy and lessons from the heart. *Nature* **453**, 322–329 (2008).
- Smith, R. R. *et al.* Regenerative potential of cardiomyocyte-derived cells expanded from percutaneous endomyocardial biopsy specimens. *Circulation* **115**, 896–908 (2007).
- Pfister, O. *et al.* Role of the ATP-binding cassette transporter Abcg2 in the phenotype and function of cardiac side population cells. *Circ. Res.* **103**, 825–835 (2008).
- Tallini, Y. N. *et al.* c-kit expression identifies cardiovascular precursors in the neonatal heart. *Proc. Natl Acad. Sci. USA* **106**, 1808–1813 (2009).
- Kubo, H. *et al.* Increased cardiac myocyte progenitors in failing human hearts. *Circulation* **118**, 649–657 (2008).
- Bearzi, C. *et al.* Human cardiac stem cells. *Proc. Natl Acad. Sci. USA* **104**, 14068–14073 (2007).
- This study describes the isolation of c-KIT⁺ CPCs from human hearts, and reports their differentiation into cardiomyocytes and vascular elements both *in vitro* and after transplantation.**
- Tang, X. L. *et al.* Intracoronary administration of cardiac progenitor cells alleviates left ventricular dysfunction in rats with a 30-day-old infarction. *Circulation* **121**, 293–305 (2010).
- Zaruba, M. M., Soonpaa, M., Reuter, S. & Field, L. J. Cardiomyogenic potential of c-kit⁺-expressing cells derived from neonatal and adult mouse hearts. *Circulation* **121**, 1992–2000 (2010).
- Patella, V. *et al.* Stem cell factor in mast cells and increased mast cell density in idiopathic and ischemic cardiomyopathy. *Circulation* **97**, 971–978 (1998).
- Sandstedt, J., Jonsson, M., Lindahl, A., Jeppsson, A. & Asp, J. C-kit⁺ CD45⁺ cells found in the adult human heart represent a population of endothelial progenitor cells. *Basic Res. Cardiol.* **105**, 545–556 (2010).
- Messina, E. *et al.* Isolation and expansion of adult cardiac stem cells from human and murine heart. *Circ. Res.* **95**, 911–921 (2004).
- Chimenti, I. *et al.* Relative roles of direct regeneration versus paracrine effects of human cardiomyocyte-derived cells transplanted into infarcted mice. *Circ. Res.* **106**, 971–980 (2010).
- Andersen, D. C., Andersen, P., Schneider, M., Jensen, H. B. & Sheikh, S. P. Murine ‘cardiospheres’ are not a source of stem cells with cardiomyogenic potential. *Stem Cells* **27**, 1571–1581 (2009).
- Orlic, D. *et al.* Bone marrow cells regenerate infarcted myocardium. *Nature* **401**, 701–705 (2001).

53. Murry, C. E. *et al.* Haematopoietic stem cells do not transdifferentiate into cardiac myocytes in myocardial infarcts. *Nature* **428**, 664–668 (2004).
54. Balsam, L. B. *et al.* Haematopoietic stem cells adopt mature haematopoietic fates in ischaemic myocardium. *Nature* **428**, 668–673 (2004).
55. Toma, C., Pittenger, M. F., Cahill, K. S., Byrne, B. J. & Kessler, P. D. Human mesenchymal stem cells differentiate to a cardiomyocyte phenotype in the adult murine heart. *Circulation* **105**, 93–98 (2002).
56. Mirosou, M. *et al.* Secreted frizzled related protein 2 (Sfrp2) is the key Akt-mesenchymal stem cell-released paracrine factor mediating myocardial survival and repair. *Proc. Natl Acad. Sci. USA* **104**, 1643–1648 (2007).
57. Hatzistergos, K. E. *et al.* Bone marrow mesenchymal stem cells stimulate cardiac stem cell proliferation and differentiation. *Circ. Res.* **107**, 913–922 (2010).
58. Chen, S. L. *et al.* Effect on left ventricular function of intracoronary transplantation of autologous bone marrow mesenchymal stem cell in patients with acute myocardial infarction. *Am. J. Cardiol.* **94**, 92–95 (2004).
59. Hare, J. M. *et al.* A randomized, double-blind, placebo-controlled, dose-escalation study of intravenous adult human mesenchymal stem cells (prochymal) after acute myocardial infarction. *J. Am. Coll. Cardiol.* **54**, 2277–2286 (2009).
60. Kehat, I. *et al.* Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *J. Clin. Invest.* **108**, 407–414 (2001).
61. Sartiani, L. *et al.* Developmental changes in cardiomyocytes differentiated from human embryonic stem cells: a molecular and electrophysiological approach. *Stem Cells* **25**, 1136–1144 (2007).
62. He, J. Q., Ma, Y., Lee, Y., Thomson, J. A. & Kamp, T. J. Human embryonic stem cells develop into multiple types of cardiac myocytes: action potential characterization. *Circ. Res.* **93**, 32–39 (2003).
63. Zhu, W. Z., Santana, L. F. & Laflamme, M. A. Local control of excitation-contraction coupling in human embryonic stem cell-derived cardiomyocytes. *PLoS ONE* **4**, e5407 (2009).
64. Zhang, J. *et al.* Functional cardiomyocytes derived from human induced pluripotent stem cells. *Circ. Res.* **104**, e30–e41 (2009).
65. Zwi, L. *et al.* Cardiomyocyte differentiation of human induced pluripotent stem cells. *Circulation* **120**, 1513–1523 (2009).
66. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
67. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
68. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
69. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
70. Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**, 618–630 (2010).
71. Shi, Y. *et al.* Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* **3**, 568–574 (2008).
72. Nussbaum, J. *et al.* Transplantation of undifferentiated murine embryonic stem cells in the heart: teratoma formation and immune response. *FASEB J.* **21**, 1345–1357 (2007).
73. Kehat, I. *et al.* Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *J. Clin. Invest.* **108**, 407–414 (2001).
74. Laflamme, M. A. *et al.* Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nature Biotechnol.* **25**, 1015–1024 (2007).
- This study describes methods for guiding the differentiation of human ESCs into cardiomyocytes, and shows that transplantation of these cells improves function in a rat infarct model.**
75. Yang, L. *et al.* Human cardiovascular progenitor cells develop from a KDR⁺ embryonic-stem-cell-derived population. *Nature* **453**, 524–528 (2008).
- This paper describes the guided differentiation and isolation of multipotent cardiovascular progenitor cells from human ESCs.**
76. Xu, X. Q. *et al.* Chemically defined medium supporting cardiomyocyte differentiation of human embryonic stem cells. *Differentiation* **76**, 958–970 (2008).
77. Zhu, W. Z. *et al.* Neuregulin/ErbB signaling regulates cardiac subtype specification in differentiating human embryonic stem cells. *Circ. Res.* **107**, 776–786 (2010).
78. Paige, S. L. *et al.* Endogenous Wnt/ β -catenin signaling is required for cardiac differentiation in human embryonic stem cells. *PLoS ONE* **5**, e11134 (2010).
79. Ueno, S. *et al.* Biphasic role for Wnt/ β -catenin signaling in cardiac specification in zebrafish and embryonic stem cells. *Proc. Natl Acad. Sci. USA* **104**, 9685–9690 (2007).
80. Bondu, A. *et al.* Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. *Cell Stem Cell* **3**, 69–84 (2008).
81. Wu, S. M. *et al.* Developmental origin of a bipotential myocardial and smooth muscle cell precursor in the mammalian heart. *Cell* **127**, 1137–1150 (2006).
82. Moretti, A. *et al.* Multipotent embryonic *Isl1*⁺ progenitor cells lead to cardiac, smooth muscle, and endothelial cell diversification. *Cell* **127**, 1151–1165 (2006).
83. Domian, I. J. *et al.* Generation of functional ventricular heart muscle from mouse ventricular progenitor cells. *Science* **326**, 426–429 (2009).
84. Caspi, O. *et al.* Transplantation of human embryonic stem cell-derived cardiomyocytes improves myocardial performance in infarcted rat hearts. *J. Am. Coll. Cardiol.* **50**, 1884–1893 (2007).
85. Fernandes, S. *et al.* Human embryonic stem cell-derived cardiomyocytes engraft but do not alter cardiac remodeling after chronic infarction in rats. *J. Mol. Cell Cardiol.* **49**, 941–949 (2010).
86. van Laake, L. W., Passier, R., Doevendans, P. A. & Mummery, C. L. Human embryonic stem cell-derived cardiomyocytes and cardiac repair in rodents. *Circ. Res.* **102**, 1008–1010 (2008).
87. Murry, C. E., Kay, M. A., Bartosek, T., Hauschka, S. D. & Schwartz, S. M. Muscle differentiation during repair of myocardial necrosis in rats via gene transfer with MyoD. *J. Clin. Invest.* **98**, 2209–2217 (1996).
88. Ieda, M. *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386 (2010).
- This study describes the direct reprogramming of cardiac fibroblasts to a cardiomyocyte-like state by the forced expression of three cardiac transcription factors.**
89. Efe, J. A. *et al.* Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. *Nature Cell Biol.* **13**, 215–222 (2011).
90. Ratner, B. D. & Bryant, S. J. Biomaterials: where we have been and where we are going. *Annu. Rev. Biomed. Eng.* **6**, 41–75 (2004).
91. Madden, L. R. *et al.* Proangiogenic scaffolds as functional templates for cardiac tissue engineering. *Proc. Natl Acad. Sci. USA* **107**, 15211–15216 (2010).
92. Radisic, M. *et al.* Functional assembly of engineered myocardium by electrical stimulation of cardiac myocytes cultured on scaffolds. *Proc. Natl Acad. Sci. USA* **101**, 18129–18134 (2004).
93. Zimmermann, W. H. *et al.* Engineered heart tissue grafts improve systolic and diastolic function in infarcted rat hearts. *Nature Med.* **12**, 452–458 (2006).
- This paper describes the formation of engineered heart tissue using neonatal rat cardiomyocytes, and shows that their implantation improves electrical and mechanical function in injured hearts.**
94. Reinecke, H., Zhang, M., Bartosek, T. & Murry, C. E. Survival, integration, and differentiation of cardiomyocyte grafts: a study in normal and injured rat hearts. *Circulation* **100**, 193–202 (1999).
95. Stevens, K. R. *et al.* Physiological function and transplantation of scaffold-free and vascularized human cardiac muscle tissue. *Proc. Natl Acad. Sci. USA* **106**, 16568–16573 (2009).
96. Caspi, O. *et al.* Tissue engineering of vascularized cardiac muscle from human embryonic stem cells. *Circ. Res.* **100**, 263–272 (2007).
97. Dvir, T. *et al.* Prevascularization of cardiac patch on the omentum improves its therapeutic outcome. *Proc. Natl Acad. Sci. USA* **106**, 14990–14995 (2009).
98. Zakharova, L. *et al.* Transplantation of cardiac progenitor cell sheet onto infarcted heart promotes cardiogenesis and improves function. *Cardiovasc Res.* **87**, 40–49 (2010).
99. Brodsky, V. Y., Chernyaev, A. L. & Vasilyeva, I. A. Variability of the cardiomyocyte ploidy in normal human hearts. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **61**, 289–294 (1991).
100. Rumyantsev, P. P., Erokhina, I. L., Antipanova, E. M. & Martynova, M. G. DNA and sex chromatin content in nuclei of conductive system and working myocytes of normal and hypertrophied human heart. *Acta Histochem. Suppl.* **39**, 225–237 (1990).

Acknowledgements We thank our many colleagues whose useful discussions and work contributed to the material presented here. This work was supported by grants from the US National Institutes of Health.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to C.E.M. (murry@u.washington.edu).

Global quantification of mammalian gene expression control

Björn Schwanhäusser¹, Dorothea Busse¹, Na Li¹, Gunnar Dittmar¹, Johannes Schuchhardt², Jana Wolf¹, Wei Chen¹ & Matthias Selbach¹

Gene expression is a multistep process that involves the transcription, translation and turnover of messenger RNAs and proteins. Although it is one of the most fundamental processes of life, the entire cascade has never been quantified on a genome-wide scale. Here we simultaneously measured absolute mRNA and protein abundance and turnover by parallel metabolic pulse labelling for more than 5,000 genes in mammalian cells. Whereas mRNA and protein levels correlated better than previously thought, corresponding half-lives showed no correlation. Using a quantitative model we have obtained the first genome-scale prediction of synthesis rates of mRNAs and proteins. We find that the cellular abundance of proteins is predominantly controlled at the level of translation. Genes with similar combinations of mRNA and protein stability shared functional properties, indicating that half-lives evolved under energetic and dynamic constraints. Quantitative information about all stages of gene expression provides a rich resource and helps to provide a greater understanding of the underlying design principles.

The four fundamental cellular processes involved in gene expression are transcription, mRNA degradation, translation and protein degradation. It is now clear that each step of this cascade is controlled by gene-regulatory events^{1,2}. Although each individual process has been intensively studied, little is known about how the combined effect of all regulatory events shapes gene expression. The fundamental question of how genomic information is processed at different levels to obtain a specific cellular proteome has therefore remained unanswered.

With regard to a quantitative description of gene expression, numerous previous studies comparing mRNA and protein levels concluded that the correlation is poor^{3,4}. However, the available data suffer from several limitations. Most studies are limited to a few hundred genes, mainly due to the technical challenges involved in large-scale protein identification and quantification. Also, protein levels measured in one experiment are typically compared to mRNA levels determined in a different experiment performed at a different time in a different laboratory, making it difficult to interpret why the correlation is low. Finally, mRNA and protein levels result from coupled processes of synthesis and degradation. Therefore, analysis of mRNA and protein levels alone cannot provide sufficient information to understand gene expression comprehensively. mRNA and protein turnover can be measured with drugs to inhibit transcription or translation^{5,6}, but this has severe side effects. Studies based on artificial fusion proteins are problematic because tagging can affect protein stability⁷.

To overcome these limitations we sought to quantify cellular mRNA and protein expression levels and turnover in parallel in a population of unperturbed mammalian cells. Pulse labelling with radioactive nucleosides or amino acids is regarded as the gold standard method to determine mRNA and protein half-lives. Recently, variants of this approach based on non-radioactive tracers have been established^{8–10}. In stable isotope labelling by amino acids in cell culture (SILAC), cells are cultivated in a medium containing heavy stable-isotope versions of essential amino acids¹¹. When non-labelled (that is, light) cells are transferred to heavy SILAC growth medium, newly synthesized proteins incorporate the heavy label while pre-existing proteins remain in the

light form. This strategy can be used to measure protein turnover^{12–14} or relative changes in protein translation^{15,16}. Similarly, newly synthesized RNA can be labelled with the nucleoside analogue 4-thiouridine (4sU). 4sU-containing mRNA can be purified and compared with the pre-existing fraction to compute mRNA half-lives¹⁰.

Pulse labelling of proteins and mRNAs

We used parallel metabolic pulse labelling with amino acids and 4sU to measure simultaneously protein and mRNA turnover in a population of exponentially growing non-synchronized NIH3T3 mouse fibroblasts (Fig. 1a). Protein samples were collected at three time points, measured by liquid chromatography and online tandem mass spectrometry (LC-MS/MS) and analysed with the MaxQuant software package¹⁷. We identified 84,676 peptide sequences and assigned them to 6,445 unique proteins (false discovery rate <1% at the peptide and protein level). A total of 5,279 of these proteins was quantified by at least three heavy to light (H/L) peptide ratios (Fig. 1b). Tissue-specific amino acid precursor pools and recycling rates, a pervasive problem for *in vivo* pulse labelling experiments^{9,18,19}, did not appreciably affect our results (Supplementary Fig. 1). For constant incorporation rates the logarithm of H/L ratios should increase linearly with time (Fig. 1c). Ninety-three per cent of proteins showed excellent linear correlation indicated by a variability of the linear regression slope smaller than 1% (Fig. 1d). Protein abundance did not influence H/L ratio measurements (Supplementary Fig. 2). In total, we obtained a confident set of 5,028 protein half-lives calculated from the slope of the regression line. Cycloheximide chase experiments for selected proteins spanning a representative range of half-lives agreed well with half-lives determined by pulsed labelling and mass spectrometry (Supplementary Fig. 3). In parallel, we pulse labelled newly synthesized RNA for 2 h with 4sU. RNA samples were fractionated into the newly synthesized and pre-existing fractions. Both fractions and the total RNA sample were analysed by mRNA sequencing and quantified by mapping reads to their exonic region²⁰. We calculated mRNA half-lives based on the ratios of newly synthesized RNA/total RNA ratio and the pre-existing RNA/total RNA¹⁰.

¹Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, D-13092 Berlin, Germany. ²MicroDiscovery GmbH, Marienburger Str. 1, D-10405 Berlin, Germany.

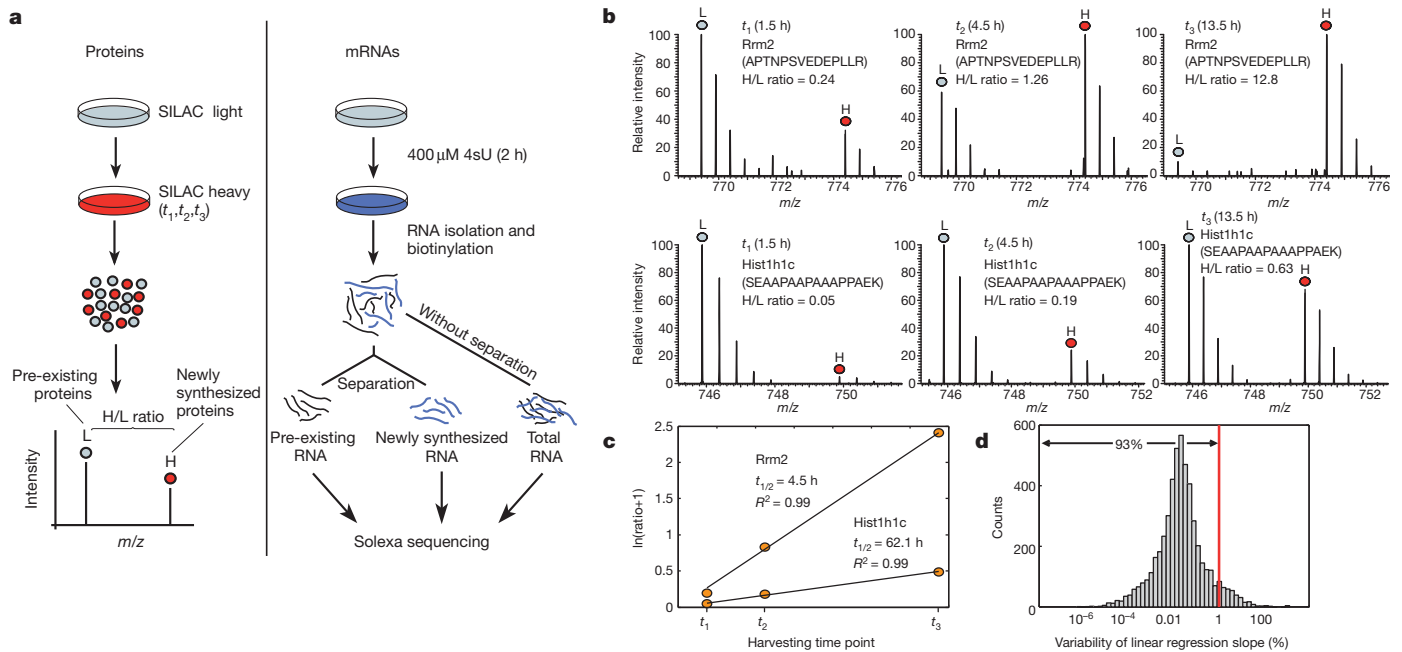


Figure 1 | Parallel quantification of mRNA and protein turnover and levels. **a**, Mouse fibroblasts were pulse labeled with heavy amino acids (SILAC, left) and the nucleoside 4-thiouridine (4sU, right). Protein and mRNA turnover was quantified by mass spectrometry and next-generation sequencing, respectively. **b**, Mass spectra of peptides from a high- and low-turnover protein reveal

Proteins were, on average, five times more stable (median half-life of 46 h) than mRNAs (9 h) and spanned a bigger dynamic range (Fig. 2a). Because very long (>200 h) and very short (<30 min) protein half-lives cannot be accurately quantified from our three time points, the true dynamic range of protein stabilities may be even higher. Notably, we found no correlation between protein and mRNA half-lives (Fig. 2c, $R^2 = 0.02$, log-log scale).

Absolute mRNA and protein copy numbers

We calculated absolute cellular mRNA copy numbers based on the number of sequencing reads in the unfractionated sample in conjunction

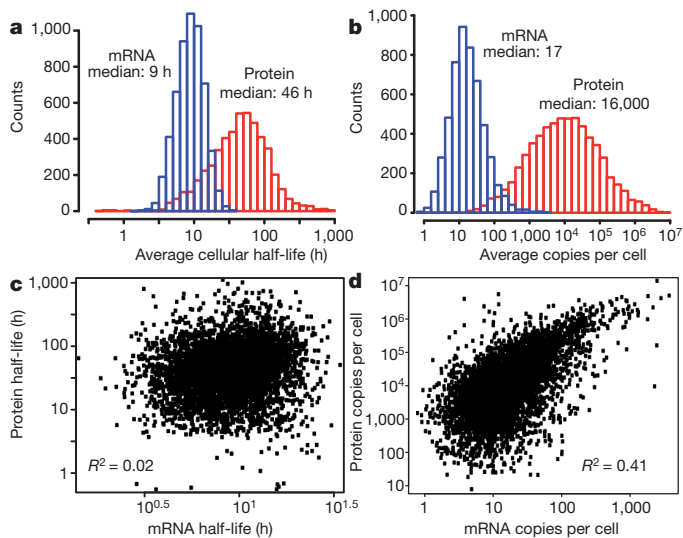


Figure 2 | mRNA and protein levels and half-lives. **a**, **b**, Histograms of mRNA (blue) and protein (red) half-lives (**a**) and levels (**b**). Proteins were on average 5 times more stable and 900 times more abundant than mRNAs and spanned a higher dynamic range. **c**, **d**, Although mRNA and protein levels correlated significantly, correlation of half-lives was virtually absent.

increasing heavy to light (H/L) ratios over time. **c**, Protein half-lives were calculated from log H/L ratios at all three time points using linear regression. **d**, Variability of linear regression slopes assessed by leave-one-out cross-validation was small.

with information on cellular mRNA content²⁰. Absolute protein copy numbers can be inferred from mass spectrometry data^{21,22}. To this end, we used the sum of peak intensities of all peptides matching to a specific protein. When divided by the number of theoretically observable peptides, this value provides an accurate proxy for protein levels ('intensity-based absolute quantification' or iBAQ, see Supplementary Methods).

Levels of detected proteins spanned approximately five orders of magnitude (Fig. 2b). Relatively few proteins had less than 100 copies per cell, indicating that some proteins of low abundance escaped detection. Indeed, we observed a moderate detection bias (Supplementary Fig. 4) and therefore restricted our analysis to genes that were identified at both the mRNA and protein level. In this subset, proteins were, on average, ~ 900 times more abundant than corresponding transcripts. Despite a huge spread, mRNA and protein levels were clearly correlated (Fig. 2d, $R^2 = 0.41$, log-log scale). This correlation is considerably higher than in any previous study in mammals^{3,4,23}. An attempt to improve this correlation further by nonlinear transformation resulted only in a marginal increase ($R^2 = 0.44$, Supplementary Fig. 5). It seems that for our data set, this is about the maximum correlation between mRNA and protein that can be achieved without additional information.

Reproducibility

To investigate the experimental noise we performed a second independent large-scale experiment and measured mRNA and protein levels and half-lives again. The overall correlation of half-lives and levels between both replicates was good (Supplementary Fig. 6 and Supplementary Table 1). Removing less-consistent data points did not increase correlation between mRNA and protein levels or half-lives (Supplementary Fig. 7). Thus, noise has little impact on the observed correlation between mRNA and protein levels and half-lives. We also validated absolute mRNA and protein copy numbers using independent methods. For mRNA copy numbers we used the NanoString technology, which captures and counts individual transcripts without enzymatic reactions²⁴. Correlation between sequencing and NanoString data was high ($r = 0.79$, see also Supplementary Fig. 8a). Absolute protein quantification was validated by spike-in

experiments using a mixture of 48 proteins with known concentrations (Supplementary Fig. 8b). iBAQ values correlated well with known absolute protein amounts over at least four orders of magnitude and had a higher precision and accuracy than alternative measures of absolute protein abundance (data not shown)^{21,22}. We also assessed degradation and synthesis rates for mRNAs and proteins by actinomycin D and cycloheximide treatment, respectively. For high turnover proteins and mRNAs we obtained results consistent with pulse labelling data (Supplementary Fig. 8c–f).

A quantitative model of gene expression

Our data allow us to calculate average synthesis rates of mRNAs and proteins for thousands of genes using a mathematical model (Fig. 3a and Supplementary Methods). The experimental data are based on a population of non-synchronized cells. Therefore, our estimated rates provide an average over the population and time.

Average cellular transcription rates predicted by the model spanned two orders of magnitude with a median of about two mRNA molecules per hour (Fig. 3b). An extreme example was *Mdm2* with more than 500 mRNAs per hour. A microscopic study on the cytomegalovirus (CMV) promoter reported transcription termination rates of 5.8 to 8.7 mRNAs per hour²⁵. These values are above the median of our predictions, as perhaps expected for a strong promoter system. Next, we calculated translation rate constants; that is, how many proteins are made from each mRNA template per hour (Fig. 3c). We find a median translation rate constant of about 40 proteins per mRNA per hour. Several proteins involved in translational regulation—such as the translation initiation factor eIF4G1, fragile X syndrome related protein Fxr2 and tuberlin—had extremely low rate constants and were translationally repressed. Plotting translation rate constants against protein levels revealed that abundant proteins are translated about 100 times more efficiently than those of low abundance (Fig. 3d). Hence, different translation efficiencies contribute to the higher dynamic range of proteins compared to mRNAs (Fig. 2b). Intriguingly, translation rate constants saturated at around 180 protein copies per mRNA per hour. To our knowledge, the maximal

translation rate constant in mammals is not known. On the basis of ref. 1, the estimated maximal translation rate constant in sea urchin embryos is 140 copies per mRNA per hour, which is surprisingly close to the prediction of our model.

Control of gene expression

A long-standing question is how much protein abundance is controlled at the transcriptional, post-transcriptional, translational and post-translational levels. Until now, this has mainly been addressed indirectly by analysing mRNA and protein sequence features. Features related to translation initiation (for example, Shine–Dalgarno, Kozak and 3′ untranslated region (UTR) sequences), elongation (for example, codon bias) and protein stability (for example, degrons) have been analysed and reported to correlate partially with protein/mRNA ratios in bacteria, yeast and mammals^{23,26,27}. We also observed sequence features characteristic of mRNA and protein stability and found that mRNAs with long 3′ UTRs are, on average, less stable (Supplementary Fig. 9). In addition, the density of AU-rich elements and binding motifs of a specific RNA-binding protein (pumilio 2) correlated negatively with mRNA stability (Supplementary Fig. 10). Highly structured proteins were more stable than unstructured ones (Supplementary Fig. 11a). We also identified amino acids over-represented in unstable proteins (Supplementary Fig. 11b).

Sequence features are at best indirect proxies for mechanisms controlling protein abundance. How much efficiencies of different steps in the gene expression cascade contribute to variance of cellular protein copy numbers can only be revealed by direct parallel genome-scale measurements of mRNA and protein levels and half-lives which were not available previously. In our data the coefficient of determination (R^2) between mRNA and protein copy numbers is 0.41 (Fig. 2d). Assuming the absence of technical and biological noise, this means that ~40% of the variance in protein levels is explained by mRNA levels—considerably more than previously thought (Fig. 4a). Most of this 40% is due to different transcription rates, whereas mRNA stability has a smaller role. Considering translation rate constants markedly boosts R^2 to 0.95. Thus, translation rate constants have the dominant role for control of protein levels. Unexpectedly, the impact of protein degradation is rather small.

In the above analysis the same experimental data were used to calculate synthesis rates and to estimate their impact on protein levels. To avoid this over-fit and to assess reliability of the model predictions we performed the same analysis with data from the biological replicate experiment. In the replicate the coefficient of determination between mRNA and protein levels was 0.37 (Fig. 4b). We then used the model including the estimated parameters from the first experiment to predict protein levels from mRNA levels in the replicate data. Predicted protein levels agreed very well with measured protein levels ($R^2 = 0.85$, Fig. 4c). Therefore, the model explains ~85% of the variability in protein copy numbers in an independent experiment. The correlation is very similar to the direct comparison of protein levels in both experiments ($R^2 = 0.84$, Supplementary Fig. 6d). We conclude that technical and biological noise in our data are low, and that the model faithfully predicts protein levels from mRNA levels in mouse fibroblasts. It also indicates that the estimated impact of transcription, mRNA stability, translation and protein stability on protein abundance is reproducible. We finally assessed how much of the efficiencies of the various steps in gene expression are retained in a different cell type and organism. To this end, we quantified mRNA and protein abundance in the human breast cancer cell line MCF7 by RNA-seq and mass spectrometry, respectively. A total of 2,030 human genes from the MCF7 data set had orthologues in the mouse fibroblast data. We then used rates from the mouse fibroblast model to predict protein levels from mRNA levels in human breast cancer cells. In MCF7 cells, the model predicted ~60% of the variability in protein levels (Fig. 4a). Although the fraction explained by the model is smaller than in mouse fibroblasts, this indicates that translation and degradation rates are to

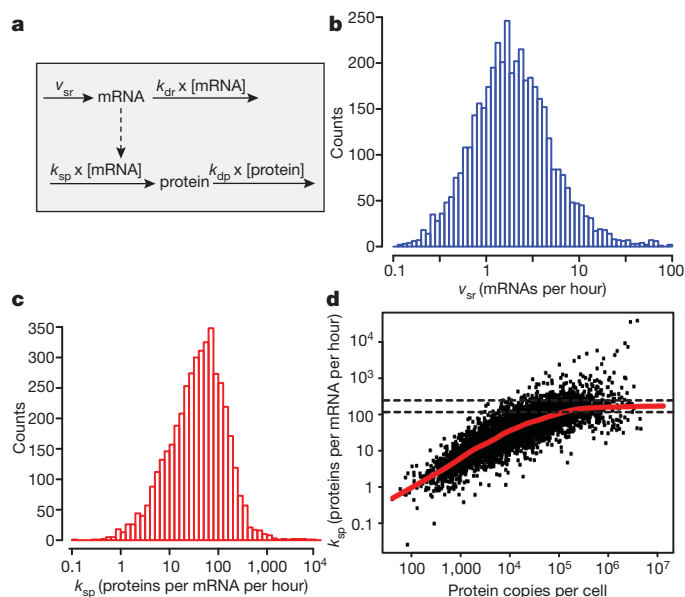


Figure 3 | Quantitative model of gene expression in growing cells.

a, mRNAs are synthesized with the rate v_{sr} and degraded with a rate constant k_{dr} . Proteins are translated and degraded with rate constants k_{sp} and k_{dp} , respectively. **b**, Calculated mRNA transcription rates show a uniform distribution. **c**, Calculated translation rate constants are not uniform. **d**, Translation rate constants of abundant proteins saturate between approximately 120 and 240 proteins per mRNA per hour. Red line shows the locally weighted fit (Lowess). Dashed lines indicate 95% confidence intervals of the Lowess maximum value calculated by bootstrapping.

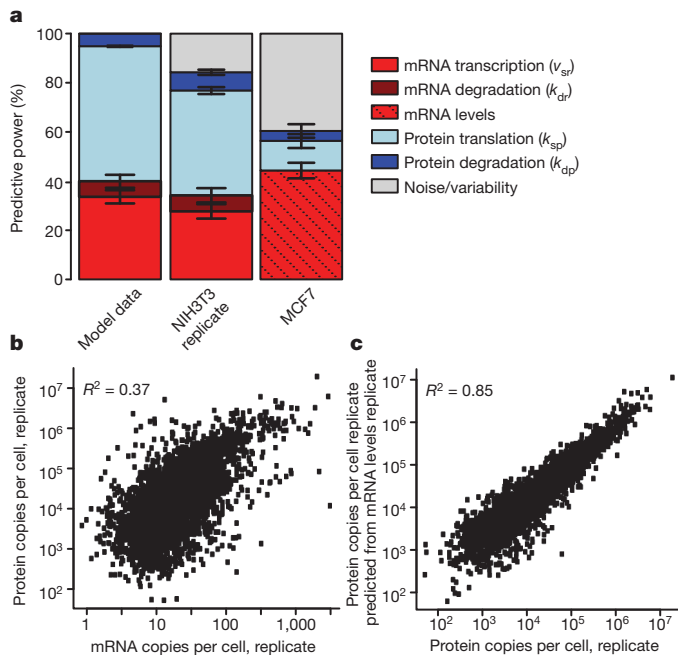


Figure 4 | Impact of different rates and rate constants on protein abundance. **a**, Protein levels are best explained by translation rates, followed by transcription rates. mRNA and protein stability is less important (left bar). **b**, In the replicate experiment mRNA levels explained 37% of protein levels in NIH3T3 cells (middle bar in **a**). **c**, The model explains 85% of variance in protein levels from measured mRNA levels (middle bar in **a**). The mouse fibroblast model has some predictive power for human orthologous genes in MCF7 cells (right bar in **a**). Error bars show 95% confidence intervals estimated by bootstrapping.

some extent independent of the cell type and conserved between mouse and human. It is noticeable, however, that the drop in prediction is mainly due to the fact that the translation part of the model performs less well.

Half-lives and gene function

Degradation of proteins is critically involved in many cellular processes including cell-cycle progression, signal transduction and apoptosis^{28–30}. Similarly, mRNA stability is important for the temporal order of gene induction^{10,31}. Genes may have evolved specific combinations of mRNA and protein half-lives under functional constraints^{10,31,32}. We therefore asked if genes with specific combinations of mRNA and protein stability have distinct biological functions. We grouped genes according to their half-lives and used gene ontology to find enriched biological processes (Fig. 5; see Supplementary Table 2 for a complete list).

Genes with stable mRNAs and stable proteins were enriched in constitutive cellular processes like translation (that is, ribosomal proteins), respiration and central metabolism (glycolysis, citric acid cycle). Hence, many housekeeping genes tend to have stable mRNAs and proteins. In yeast energy costs keep transcription and translation rates under selective pressure³³. We reasoned that energy constraints may explain why housekeeping genes tend to have stable mRNAs and proteins. On the basis of the model, we calculated the theoretical energy required to maintain cellular mRNA and protein levels by recycling from their building blocks (nucleotide monophosphates and amino acids, respectively) in terms of high energy phosphates. This is a conservative estimate as splicing, folding and transport are not included. Protein synthesis consumes more than 90% of the energy whereas less than 10% is needed for transcription. A total of 20% of the proteins consumed 80% of the energy for translation (Pareto principle or 80/20 rule). Consistent with optimization under energy constraints, abundant proteins were significantly more stable than less abundant ones (Supplementary Fig. 12a, $P < 10^{-15}$,

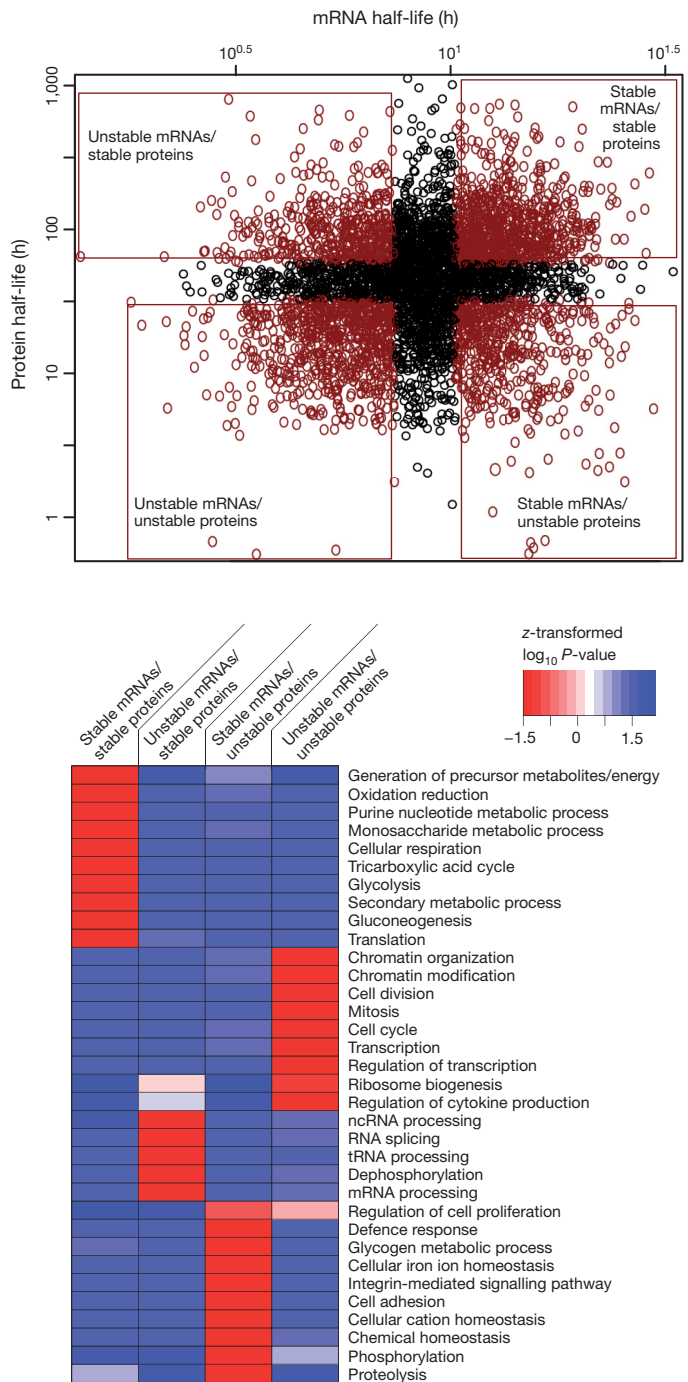


Figure 5 | Functional characteristics of genes with different mRNA and protein half-lives. Genes were grouped according to their combination of mRNA and protein half-lives and analysed for enriched gene ontology terms. A heat map of enrichment P -values reveals functional similarities of genes with similar combinations of half-lives.

Wilcoxon test). This is not necessarily expected because the overall contribution of protein stability to protein levels is very small (Fig. 4a). In addition, abundant proteins were significantly shorter (Supplementary Fig. 12b). Shuffling protein half-lives and lengths markedly increased theoretical energy consumption (Supplementary Fig. 12c). Collectively, these observations indicate that mammalian gene expression evolved under energy constraints.

The subset of genes with unstable mRNAs and proteins was strongly enriched in transcription factors, signalling genes, chromatin modifying enzymes and genes with cell-cycle-specific functions (Fig. 5). Because mRNAs and proteins are information carriers, their degradation can be

interpreted as a built-in timer that controls the persistence of genetic information³⁴. It therefore makes intuitive sense that many regulatory genes have short mRNA and protein half-lives. However, it must be stressed that population-level data cannot provide information about individual cells or molecules.

The group of genes with stable proteins but unstable mRNAs was strongly enriched in terms related to processing of mRNAs, tRNAs and non-coding RNAs. Hence, many mammalian RNA-binding proteins are stable whereas their encoding transcripts are short lived, as also found in yeast³⁵. Because many RNA-binding proteins bind their own message³⁶, this observation is indicative of a post-transcriptional negative feedback loop for RNA-binding proteins. Consistently, we found that unstable mRNAs are enriched for binding motifs of RNA-binding proteins (Supplementary Fig. 10).

Finally, the subset of genes with stable mRNAs and unstable proteins was rich in extracellular proteins. This is expected, as secreted proteins have a short cellular half-life. Additionally, this group contains proteins involved in cellular homeostasis, defence response and proteolysis. This set contains two ferritin proteins that are rapidly upregulated in response to iron³⁷. Ferritins are classic examples of translationally regulated genes. As translational regulation is not dependent on mRNA half-lives, genes with stable mRNAs can still be dynamically regulated as long as their protein half-lives are short. It is tempting to speculate that other homeostasis genes in this group are regulated at the level of translation.

Discussion

Although gene expression is one of the most fundamental processes in biology it has never been quantified comprehensively. We provide the first analysis of mRNA and protein levels, half-lives, transcription rates and translation rate constants for thousands of genes. In the future, additional methods like sequencing of nascent transcripts and ribosome profiling may further refine this picture^{38,39}.

We found that mRNA levels explain around 40% of the variability in protein levels. This fraction is higher than in previous studies on mammals^{3,4,23}. We found that in mouse fibroblasts, translation efficiency is the single best predictor of protein levels. Hence, protein abundance seems to be predominantly regulated at the ribosome, highlighting the importance of translational control^{40,41}. Whether this observation is valid in other cell types is not known. A recent study on embryonic stem cells revealed that changes in protein levels are not accompanied by changes in corresponding mRNAs⁴². It is also not clear how much translation rate constants change under different conditions. Our observation that the mouse model can to some degree predict levels of orthologous proteins in MCF7 cells suggests that translation efficiency is partially 'hard-coded' in the genome and is not subject to change.

Compared to translational control, protein stability seems to have a minor role in cellular protein abundance in our system. This is surprising as protein degradation is involved in the regulation of many cellular processes^{28–30}. From the global perspective, the dominance of translational regulation makes sense given the high energy costs associated with protein synthesis. However, it should also be stressed that our data set represents average values derived from a population of dividing, non-synchronized cells. At the single cell level, the role of protein degradation for protein abundance may be higher. Similarly, protein degradation may be more important upon perturbation.

Gene expression may follow certain design principles for optimal evolutionary fitness. Intriguingly, we found that genes with certain combinations of mRNA and protein half-lives share common functions, indicating that they evolved under similar constraints. One of these constraints may be energy efficiency³³. Consistently, we observed that the theoretical energy needed for gene expression is much lower than random. A second constraint may be the ability of genes to respond quickly to a stimulus. We find that many transcription factors and genes with cell-cycle-specific function have unstable mRNAs and

proteins, predisposing them to rapid transcriptional and/or translational regulation. In addition, genes with stable mRNAs but unstable proteins can be regulated quickly at the level of translation. These observations are consistent with the idea that many fast-responding genes have short protein and/or mRNA half-lives^{10,31,32,43}. The global picture is that most mRNAs and especially proteins are stable unless genes need to respond quickly to a stimulus. Owing to the trade-off between dynamic regulation and energy efficiency, this may be an optimal design.

Our data provide a rich resource for the scientific community that can be mined in many ways that are beyond the scope of this study (see Supplementary Table 3 for the entire data set). For example, we provide by far the largest data set on protein copy numbers, which contains valuable information for modelling of cellular processes and stoichiometry of protein complexes²². Half-lives of proteins and mRNAs can be used to search for properties of unstable mRNAs or proteins, and we provide a first analysis of characteristic sequence features (Supplementary Figs 9 and 10). Genome-scale quantitative data on absolute mRNA and protein levels and half-lives will certainly help to understand the complex relationships between thousands of genes and their products in biological systems.

Note added in proof: While this paper was in revision, another paper⁴⁴ reported that changes in mRNA levels in dendritic cells are mainly determined by transcription rates. This result is consistent with our findings in fibroblasts. Notably, mRNA half-lives reported in ref. 44 are considerably shorter (see Supplementary Information for a brief discussion).

METHODS SUMMARY

NIH3T3 cells grown in light (L) SILAC medium were simultaneously pulse-labelled with heavy (H) amino acids and 4-thiouridine (4sU). For proteome analysis, proteins were extracted, separated by SDS-polyacrylamide gel electrophoresis (PAGE), trypsin-digested and analysed by LC-MS/MS on high-resolution instruments (LTQ-Orbitrap XL and Velos, Thermo Fisher). Raw files were processed by MaxQuant (version 1.0.13.13) for peptide/protein identification and quantification. In total 3,588,163 fragment spectra led to 972,333 peptide identifications (84,676 unique peptide sequences) that were assigned to 6,445 unique proteins (false discovery rate of 1% at the peptide and protein level). Average absolute mass deviation was 0.29 parts per million (p.p.m.). Absolute protein amounts were calculated as the sum of all peptide peak intensities divided by the number of theoretically observable tryptic peptides (intensity based absolute quantification, or iBAQ). RNA was extracted and separated into newly synthesized and pre-existing fractions based on the incorporated 4sU. Total, pre-existing and newly synthesized RNA samples were processed according to an mRNA sequencing protocol (two rounds of oligo(dT) enrichment) and analysed on a Solexa GAII-X sequencing platform (36 cycles). Reads were mapped to the mouse genome reference sequence (mm9, July 2007) using SOAP2 with a maximum of two mismatches allowed. Only uniquely mapped reads were retained. For more details on data acquisition, processing, analysis and modelling see Supplementary Methods.

Received 16 November 2010; accepted 1 April 2011.

1. Ben-Tabou de-Leon, S. & Davidson, E. H. Modeling the dynamics of transcriptional gene regulatory networks for animal development. *Dev. Biol.* **325**, 317–328 (2009).
2. Komili, S. & Silver, P. A. Coupling and coordination in gene expression processes: a systems biology view. *Nature Rev. Genet.* **9**, 38–48 (2008).
3. de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M. & Vogel, C. Global signatures of protein and mRNA expression levels. *Mol. Biosyst.* **5**, 1512–1526 (2009).
4. Maier, T., Guell, M. & Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* **583**, 3966–3973 (2009).
5. Belle, A., Tanay, A., Bitincka, L., Shamir, R. & O'Shea, E. K. Quantification of protein half-lives in the budding yeast proteome. *Proc. Natl Acad. Sci. USA* **103**, 13004–13009 (2006).
6. Yang, E. *et al.* Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. *Genome Res.* **13**, 1863–1872 (2003).
7. Yen, H. C., Xu, Q., Chou, D. M., Zhao, Z. & Elledge, S. J. Global protein stability profiling in mammalian cells. *Science* **322**, 918–923 (2008).
8. Gouw, J. W., Krijgsveld, J. & Heck, A. J. Quantitative proteomics by metabolic labeling of model organisms. *Mol. Cell. Proteomics* **9**, 11–24 (2010).
9. Beynon, R. J. & Pratt, J. M. Metabolic labeling of proteins for proteomics. *Mol. Cell. Proteomics* **4**, 857–872 (2005).

10. Friedel, C. C., Dolken, L., Ruzsics, Z., Koszinowski, U. H. & Zimmer, R. Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. *Nucleic Acids Res.* **37**, e115 (2009).
11. Mann, M. Functional and quantitative proteomics using SILAC. *Nature Rev. Mol. Cell Biol.* **7**, 952–958 (2006).
12. Doherty, M. K., Hammond, D. E., Clague, M. J., Gaskell, S. J. & Beynon, R. J. Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. *J. Proteome Res.* **8**, 104–112 (2009).
13. Milner, E., Barnea, E., Beer, I. & Admon, A. The turnover kinetics of major histocompatibility complex peptides of human cancer cells. *Mol. Cell. Proteomics* **5**, 357–365 (2006).
14. Lam, Y. W., Lamond, A. I., Mann, M. & Andersen, J. S. Analysis of nucleolar protein dynamics reveals the nuclear degradation of ribosomal proteins. *Curr. Biol.* **17**, 749–760 (2007).
15. Schwanhäusser, B., Gossen, M., Dittmar, G. & Selbach, M. Global analysis of cellular protein translation by pulsed SILAC. *Proteomics* **9**, 205–209 (2009).
16. Selbach, M. *et al.* Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58–63 (2008).
17. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnol.* **26**, 1367–1372 (2008).
18. Price, J. C., Guan, S., Burlingame, A., Prusiner, S. B. & Ghaemmaghami, S. Analysis of proteome dynamics in the mouse brain. *Proc. Natl Acad. Sci. USA* **107**, 14508–14513 (2010).
19. Wu, C. C., MacCoss, M. J., Howell, K. E., Matthews, D. E. & Yates, J. R. III. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* **76**, 4951–4959 (2004).
20. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621–628 (2008).
21. Lu, P., Vogel, C., Wang, R., Yao, X. & Marcotte, E. M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnol.* **25**, 117–124 (2007).
22. Malmström, J. *et al.* Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature* **460**, 762–765 (2009).
23. Vogel, C. *et al.* Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. *Mol. Syst. Biol.* **6**, 400 (2010).
24. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnol.* **26**, 317–325 (2008).
25. Darzacq, X. *et al.* In vivo dynamics of RNA polymerase II transcription. *Nature Struct. Mol. Biol.* **14**, 796–806 (2007).
26. Arava, Y., Boas, F. E., Brown, P. O. & Herschlag, D. Dissecting eukaryotic translation and its control by ribosome density mapping. *Nucleic Acids Res.* **33**, 2421–2432 (2005).
27. Wu, G., Nie, L. & Zhang, W. Integrative analyses of posttranscriptional regulation in the yeast *Saccharomyces cerevisiae* using transcriptomic and proteomic data. *Curr. Microbiol.* **57**, 18–22 (2008).
28. Kirkpatrick, D. S., Denison, C. & Gygi, S. P. Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics. *Nature Cell Biol.* **7**, 750–757 (2005).
29. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
30. King, R. W., Deshaies, R. J., Peters, J. M. & Kirschner, M. W. How proteolysis drives the cell cycle. *Science* **274**, 1652–1659 (1996).
31. Hao, S. & Baltimore, D. The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. *Nature Immunol.* **10**, 281–288 (2009).
32. Legewie, S., Herzel, H., Westerhoff, H. V. & Bluthgen, N. Recurrent design patterns in the feedback regulation of the mammalian signalling network. *Mol. Syst. Biol.* **4**, 190 (2008).
33. Wagner, A. Energy constraints on the evolution of gene expression. *Mol. Biol. Evol.* **22**, 1365–1374 (2005).
34. Pedraza, J. M. & Paulsson, J. Effects of molecular memory and bursting on fluctuations in gene expression. *Science* **319**, 339–343 (2008).
35. Mittal, N., Roy, N., Babu, M. M. & Janga, S. C. Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks. *Proc. Natl Acad. Sci. USA* **106**, 20300–20305 (2009).
36. Hogan, D. J., Riordan, D. P., Gerber, A. P., Herschlag, D. & Brown, P. O. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol.* **6**, e255 (2008).
37. Hentze, M. W., Muckenthaler, M. U. & Andrews, N. C. Balancing acts: molecular control of mammalian iron metabolism. *Cell* **117**, 285–297 (2004).
38. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
39. Churchman, L. S. & Weissman, J. S. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**, 368–373 (2011).
40. Gebauer, F. & Hentze, M. W. Molecular mechanisms of translational control. *Nature Rev. Mol. Cell Biol.* **5**, 827–835 (2004).
41. Sonenberg, N. & Hinnebusch, A. G. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**, 731–745 (2009).
42. Lu, R. *et al.* Systems-level dynamic analyses of fate change in murine embryonic stem cells. *Nature* **462**, 358–362 (2009).
43. Rosenfeld, N., Elowitz, M. B. & Alon, U. Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.* **323**, 785–793 (2002).
44. Rabani, M. *et al.* Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nature Biotechnol.* doi:10.1038/nbt.1861 (24 April 2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank N. Rajewsky and L. Dölken for fruitful discussions and C. Sommer for technical assistance. M.S. and W.C. are supported by the Helmholtz Association, the German Ministry of Education and Research (BMBF) and the Senate of Berlin by funds aimed at establishing the Berlin Institute of Medical Systems Biology (BIMSB) (grant number 315362A). J.W. is supported by the ForSys-programme of the German Ministry of Education and Research (grant number 315289); D.B. by the Helmholtz Alliance on Systems Biology/MSBN; and N.L. by the China Scholarship Council CSC.

Author Contributions M.S. conceived, designed and supervised the experiments. B.S. performed wet-lab experiments, mass spectrometry and proteomic data analysis. D.B. and J.W. developed and employed the mathematical model. N.L. performed RNA-seq experiments. W.C. designed and supervised RNA-seq experiments. B.S., D.B., J.S., W.C. and M.S. analysed genome-wide data. G.D. helped in cycloheximide chase experiments and data analysis. B.S., D.B., J.S., J.W., W.C. and M.S. interpreted the data. M.S. wrote the manuscript.

Author Information Sequences have been deposited in the Sequence Read Archive under accession code SRA030871. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.W. (jana.wolf@mdc-berlin.de, for mathematical modelling), W.C. (wei.chen@mdc-berlin.de, for transcriptomics) or M.S. (matthias.selbach@mdc-berlin.de, for proteomics).

TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity

Kristine Williams^{1,2*}, Jesper Christensen^{1,2*}, Marianne Terndrup Pedersen^{1,2*}, Jens V. Johansen^{1,3}, Paul A. C. Cloos^{1,2}, Juri Rappsilber⁴ & Kristian Helin^{1,2}

Enzymes catalysing the methylation of the 5-position of cytosine (mC) have essential roles in regulating gene expression and maintaining cellular identity. Recently, TET1 was found to hydroxylate the methyl group of mC, converting it to 5-hydroxymethyl cytosine (hmC). Here we show that TET1 binds throughout the genome of embryonic stem cells, with the majority of binding sites located at transcription start sites (TSSs) of CpG-rich promoters and within genes. The hmC modification is found in gene bodies and in contrast to mC is also enriched at CpG-rich TSSs. We provide evidence further that TET1 has a role in transcriptional repression. TET1 binds a significant proportion of Polycomb group target genes. Furthermore, TET1 associates and colocalizes with the SIN3A co-repressor complex. We propose that TET1 fine-tunes transcription, opposes aberrant DNA methylation at CpG-rich sequences and thereby contributes to the regulation of DNA methylation fidelity.

The majority of CpGs in mammalian genomes are methylated. An exception to this is CpG islands, which are found in more than 60% of all mammalian gene promoters. These are often unmethylated and can be either transcriptionally active or inactive depending on other factors, including histone modifications and the activity of cell-type-specific transcription factors^{1–5}. In current models for gene regulation, CpG methylation in promoters leads to stable gene silencing, whereas the function of intragenic methylation might, like trimethylation of histone 3 lysine 36 (H3K36me3), repress the initiation of intragenic transcription⁶.

DNA methyltransferases are essential for embryogenesis, and the methylation pattern of the mammalian genome undergoes major changes during development. As an example, global waves of DNA demethylation and remethylation take place after fertilization, and gene-specific *de novo* methylation occurs during differentiation of embryonic stem (ES) cells^{6,7}. Importantly, patterns of DNA methylation are perturbed in human diseases such as imprinting disorders and cancer⁸. So far there is very limited knowledge regarding the mechanisms leading to DNA hypermethylation of CpG-island promoters in cancer, and how CpG-islands generally remain unmethylated in somatic cells.

Enzymes contributing to DNA demethylation could potentially provide a fidelity system for DNA methylation, but such enzymes were not known until recently. In a ground-breaking paper, TET1 was shown to catalyse the hydroxylation of mC⁹, which has led to the proposal of several models for how TET1 and hmC may contribute to DNA demethylation and gene regulation. One possibility is that hydroxylation of mC by TET1 might interfere with DNMT1 activity, leading to a subsequent passive loss of methylation following replication. Alternatively, hmC may be converted to cytosine through hitherto unknown enzymatic mechanisms. In addition, hydroxylation of mC may promote transcriptional de-repression by dissociation of mC-binding proteins and/or recruitment of effector proteins. The demonstration that hmC is highly abundant in ES cells and in neuronal Purkinje cells indicates that this modification is stably

present in the mammalian genome and that it might be important for gene regulation^{9,10}.

TET1 binds CpG-rich transcription start sites

TET1 is highly expressed in mouse ES cells and is rapidly downregulated during their differentiation^{9,11}. To obtain more information regarding the function of TET1, we inhibited TET1 expression in mouse ES cells using two different shRNA constructs (Fig. 1a and Supplementary Fig. 1a). The efficient knockdown of *Tet1* did not lead to any change in proliferation rate or expression of NANOG and OCT4 (Fig. 1a and Supplementary Fig. 1a, b). These data are in agreement with a recently published study¹², but in contrast to results reported by others¹³. We also observed inhibition of growth and decreased levels of NANOG in mouse ES cells when using the *Tet1* shRNA sequences published in the latter study (Supplementary Fig. 1c, d). However, as these shRNA sequences do not lead to greater knockdown efficiency than the ones we have used (Supplementary Fig. 1c), it is possible that shRNA off-target effects could cause the observed phenotype.

We determined the genome-wide location of TET1 by using two different antibodies to TET1 (Tet1-N and Tet1-C) for chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq). These experiments were performed in control or TET1-depleted mouse ES cells. The two TET1 antibodies were highly specific as shown in the examples provided in Fig. 1b and by the fact that 97–99% of the identified TET1 binding sites were not found in the TET1-depleted cells (Supplementary Fig. 2a). The majority of TET1 binding sites were found in gene bodies, with the highest density around TSSs (Fig. 1c). Gene annotation of TET1 binding sites, using a false discovery rate (FDR) < 0.01, showed that TET1 binds in the vicinity of the TSS of 6,573 genes (Fig. 1d and Supplementary Table 1), of which all tested so far have been independently validated by ChIP followed by real-time quantitative PCR (ChIP-qPCR, Supplementary Fig. 2b and data not shown). Peak detection analysis using FDR < 0.1 indicates that TET1 could have up to 9,241 target genes (Supplementary Fig. 3a). Gene

¹Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark. ²Centre for Epigenetics, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark. ³The Bioinformatics Centre, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark. ⁴Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, UK.

*These authors contributed equally to this work.

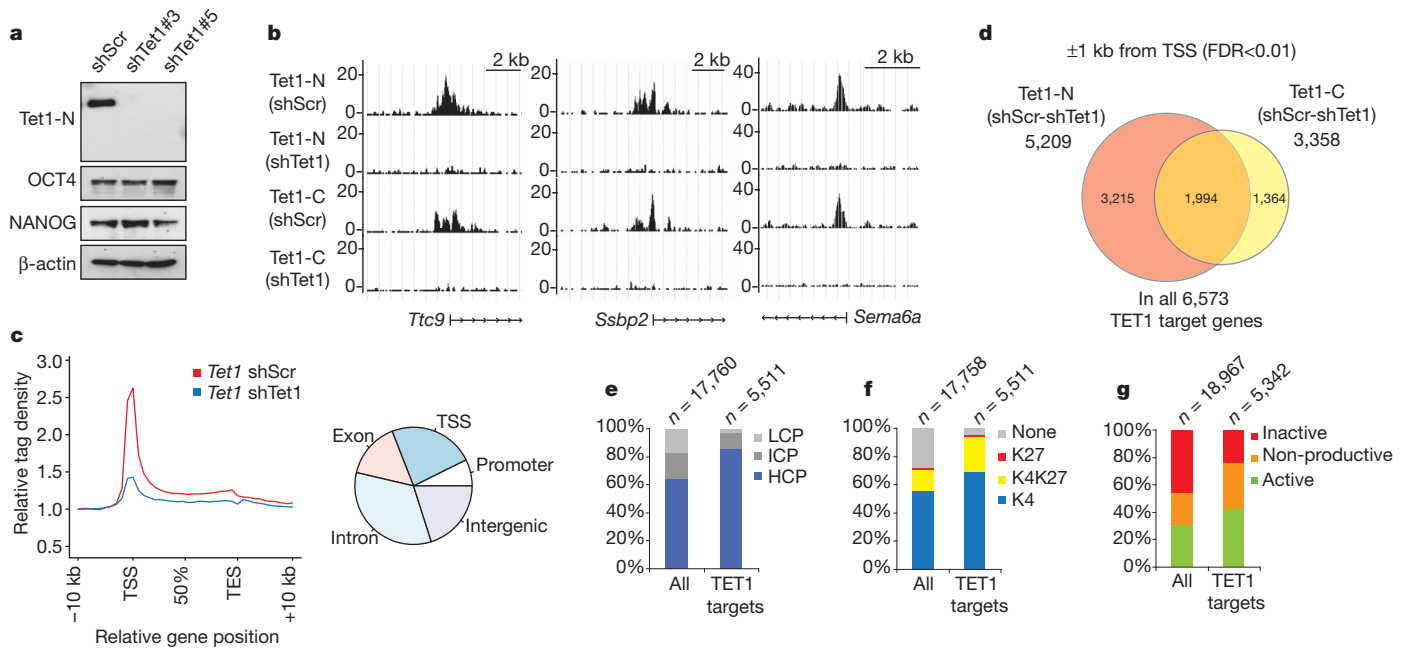


Figure 1 | Identification of TET1 target genes. **a**, Western blot showing TET1, OCT4 and NANOG levels for control-transfected (shScr) and TET1-depleted (shTet1#3 and shTet1#5) mouse ES cells. **b**, Examples of TET1 ChIP-seq results in control or *Tet1* knockdown ES cells. ChIP-seq was performed using both an anti-N- and anti-C-terminal TET1 antibody (Tet1-N and Tet1-C). y -axis of binding profiles denotes number of sequence tag reads. **c**, Left panel, mean distribution of tags across gene bodies for TET1 ChIP-seq in control and *Tet1* knockdown cells. Right panel, diagram illustrating the overall distribution of TET1 binding sites into TSS (± 1 kb), promoter (-1 to -5 kb),

exon, intron and intergenic regions. **d**, Venn diagram illustrating the overlap of TET1 target genes using anti-TET1-N and -C antibodies. **e**, **f**, Histograms showing promoter CpG density, divided into high-, intermediate- or low-density CpG promoters (HCP, ICP or LCP) as defined in ref. 25 (**e**) or distribution of H3K4me3 (K4) and H3K27me3 (K27)²⁵ (**f**) for all genes or for TET1 target genes. **g**, Overlay of TET1 target genes with active genes (RNA polymerase II binding and H3K79me2), non-productive (RNA polymerase II binding, no H3K79me2) and inactive (no RNA polymerase II binding or H3K79me2)^{26,27}.

Ontology analysis showed that TET1 target genes are involved in a variety of basic cellular processes, and in more specific processes such as development and differentiation (Supplementary Fig. 3b). The majority of the TET1 target genes are associated with high and intermediate density CpG promoters (HCPs and ICPs, Fig. 1e), which are positive for H3K4me3 (Fig. 1f). The correlation between TET1 binding and high CpG density is also found outside of TSSs (Supplementary Fig. 4). Interestingly, TET1 binding does not predict whether a promoter is active, poised for activation (non-productive) or inactive (Fig. 1g). In agreement with this, we found that a significant fraction of TET1 was associated with promoters containing the H3K27me3 Polycomb repressive mark (Fig. 1f). Indeed, independent analysis showed a highly significant overlap of genes bound by TET1 and the Polycomb group (PcG) protein, SUZ12, in ES cells (Supplementary Fig. 5a, b).

hmC is enriched at TSSs and gene bodies

To gain information regarding a possible function of hmC, we generated an affinity-purified polyclonal antibody to hmC that binds with high specificity and sensitivity to this mark, as shown by enzyme-linked immunosorbent (ELISA) and DNA immunoprecipitation (DIP) assays (Supplementary Fig. 6). Genome-wide DIP-seq assays were performed using anti-hmC, anti-mC and IgG on genomic DNA purified from control or TET1-depleted ES cells as well as from *Dnmt* triple knockout (TKO) mouse ES cells, lacking *Dnmt1*, *Dnmt3a* and *Dnmt3b*¹⁴. We confirmed by ChIP-qPCR that TET1 localizes to its target genes in the *Dnmt* TKO cells (Supplementary Fig. 7a). The analyses showed that hmC is located as discrete peaks throughout the genome (Fig. 2a). Furthermore, the majority of signals obtained with the hmC antibody were absent in *Dnmt* TKO mouse ES cells, confirming that generation of hmC requires the pre-existence of mC (Fig. 2a). The hmC modification in mouse ES cells is particularly enriched within gene bodies as also observed for the mC mark¹⁵

and recently reported for hmC in mouse cerebellum¹⁶ (Fig. 2b, c). Strikingly, in contrast to the localization of mC, hmC is also significantly enriched at the TSS coinciding with TET1 (Fig. 2c), indicating that a significant fraction of mC is converted to hmC at the TSS. Also, the hmC modification is generally not detectable at repetitive elements such as intracisternal A particle (IAP) elements and minor satellite repeats by DIP-qPCR (Supplementary Fig. 7b), further demonstrating that hmC and mC show distinct genomic distributions.

Gene annotation of hmC positive regions around the TSS (-0.7 kilobases to $+0.3$ kb) showed that 2,424 regions are hmC-positive in wild-type ES cells compared to *Dnmt* TKO ES cells. Approximately 28% of these regions showed a more than twofold reduction in hmC signal in the DIP-seq analyses upon downregulation of TET1 (Fig. 2d) and in validation experiments the knockdown of *Tet1* led to a significant decrease in hmC levels on tested genes (Fig. 2e and data not shown). Depending on the used false discovery rate cut-off for TET1, between 35% (FDR < 0.01) and 50% (FDR < 0.1) of hmC-positive genes are bound by TET1 (Fig. 2f). These results are in agreement with reports showing that *Tet1* knockdown only causes a partial decrease in global hmC levels in mouse ES cells^{9,12}, and imply that, although TET1 is important for the generation of hmC, other enzymes such as TET2 are also likely to contribute to hmC levels in mouse ES cells.

As for TET1, Gene Ontology analysis of the hmC-positive genes showed enrichment for genes involved in basic cellular processes, but also in the regulation of development and differentiation (Supplementary Fig. 7c). Moreover, hmC positivity does not correlate with transcriptional activation and surprisingly, most hmC-positive genes seem not to be expressed in mouse ES cells (Fig. 2g).

A significant proportion of the TSSs classified as positive for hmC has intermediate or high CpG content (Fig. 2h and Supplementary Fig. 4). Genome-wide analyses of the hmC distribution relative to CpG content showed that the hmC mark is enriched in regions with

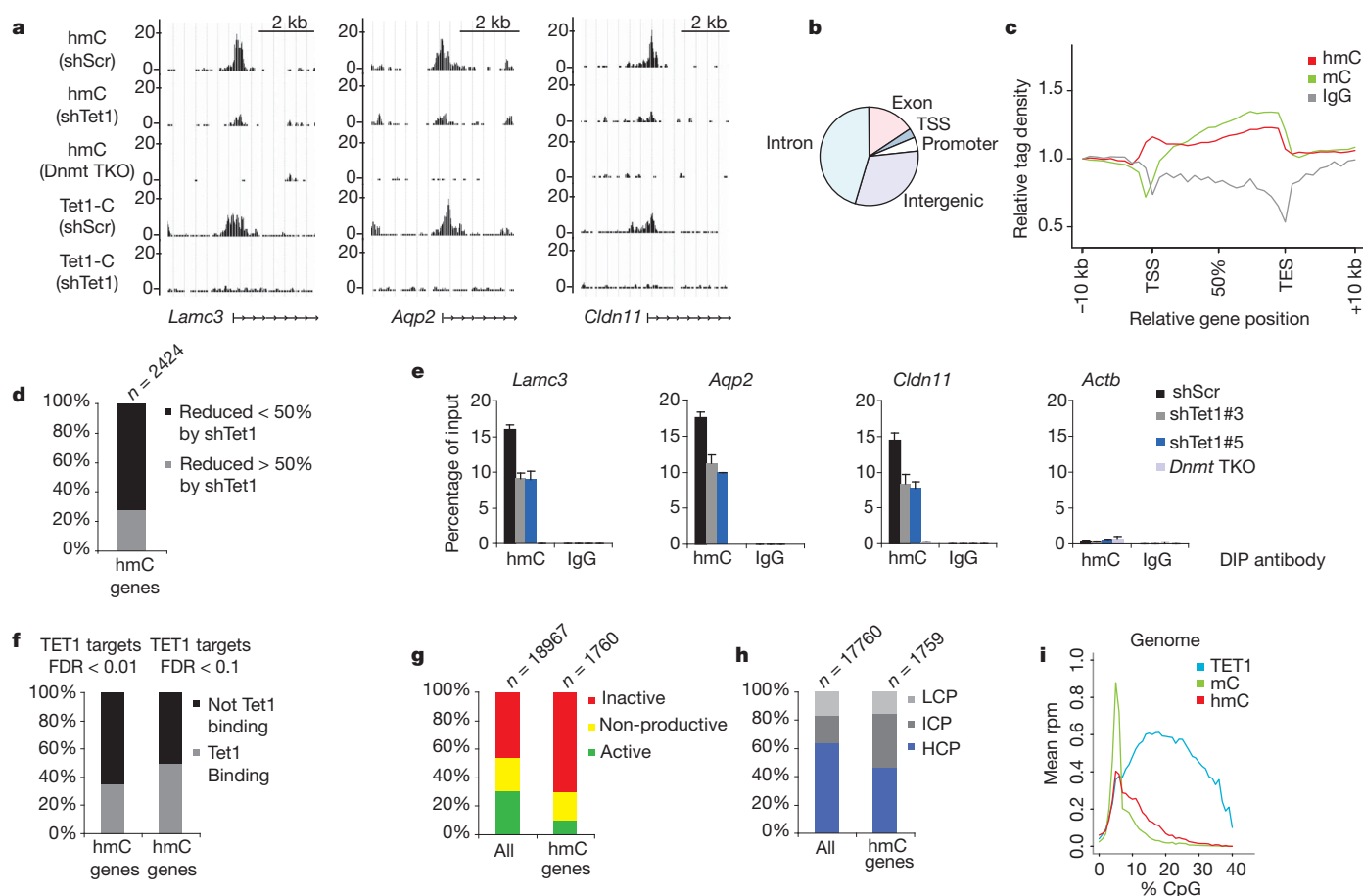


Figure 2 | Hydroxymethylcytosine localizes to TSS and gene body.

a, Examples of hmC DIP-seq results in mouse ES cells. ChIP-seq profiles of TET1 are included for comparison. **b**, Diagram illustrating the overall distribution of hmC into TSS (± 1 kb), promoter (-1 to -5 kb), exon, intron and intergenic regions. **c**, The mean distribution of tags across gene bodies for hmC, mC and IgG. **d**, Almost a third (28%) of hmC positive TSSs showed a more than twofold reduction in hmC signal in mouse ES cells depleted of TET1. **e**, DIP-qPCR was performed in control mouse ES cells, *Tet1* knockdown cells (shTet1#3 and shTet1#5), and *Dnmt* TKO cells as indicated. **f**, Overlay of genes

positive for hmC at the TSS with TET1 target genes using FDR cut-off values of 0.01 or 0.1 in the ChIP-seq analysis. **g**, Overlay of hmC positive genes with active genes (RNA polymerase II binding and H3K79me2), non-productive (RNA polymerase II binding, no H3K79me2) and inactive (no RNA polymerase II binding or H3K79me2)^{26,27}. **h**, Distribution of high-, intermediate- or low density CpG promoters (HCP, ICP or LCP)²⁵ for all genes or hmC-positive genes. **i**, Plot illustrating the genome-wide correlation of TET1, hmC and mC signal intensity (rpm, reads per million) with CpG density. All error bars denote s.d., $n = 3$.

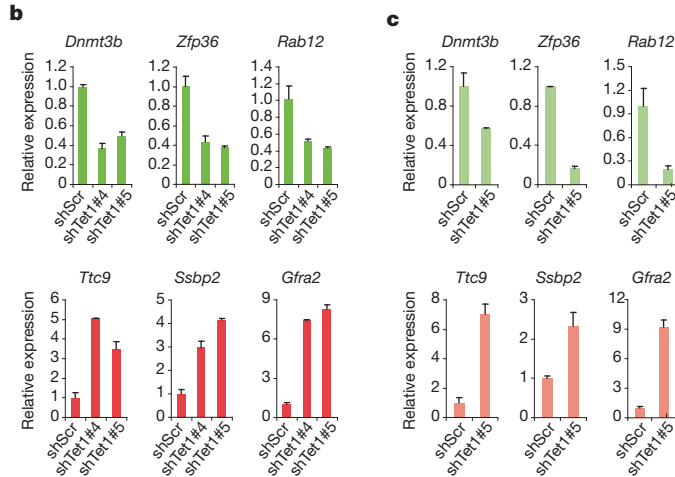
relatively high CpG content compared to mC (Fig. 2i). Whereas only 15% of hmC-positive TSSs also contain a high mC signal, we find that several hmC-positive regions have low levels of mC, implying that the two marks often co-exist. Upon *Tet1* knockdown only a minor global increase in mC was observed as evaluated by genome-wide anti-mC DIP (Me-DIP) (Supplementary Fig. 8a). However, a few hundred genes show modest TSS specific increases in mC levels after *Tet1* knockdown (Supplementary Fig. 8b). Gene Ontology analyses for these genes showed enrichments for specialized developmental processes (Supplementary Fig. 8c). Interestingly, we found that approximately a third of the genes reported to acquire DNA methylation during ES cell differentiation^{2,3} are marked by hmC in the ES cell state (Supplementary Table 2). Taken together, these results show that hmC colocalizes with mC in gene-bodies, and that hmC, in contrast to mC, is enriched at TSSs with intermediate to high CpG density, where it may contribute to the regulation of DNA methylation patterns.

TET1 contributes to transcriptional repression

To understand how TET1 contributes to the regulation of target genes, we performed genome-wide expression analyses of mouse ES cells expressing two different *Tet1* shRNAs or a scrambled shRNA (Supplementary Fig. 9a, b and Supplementary Table 3). As shown in Fig. 3a and Supplementary Fig. 9c, we observed a significant decrease

in expression of 556 genes and a significant increase in expression of 851 genes common to both shRNAs. Of these approximately 700 were direct target genes of TET1, and therefore only around 10% of all TET1 target genes change expression following *Tet1* knockdown. Whereas we expected to observe a significant fraction of the down-regulated genes to be direct targets for TET1, we were surprised to find that an even higher fraction of the upregulated genes were associated with TET1 (Fig. 3a). To validate these results, we performed qPCR analysis of a number of downregulated and upregulated genes (Fig. 3b) that were also directly bound by TET1 (Supplementary Fig. 2b). Moreover, several of the identified targets show similar expression change upon differentiation of mouse ES cells by retinoic acid, which leads to decreased levels of TET1 (Supplementary Fig. 9d).

To investigate whether the transcriptional effects of TET1 are mediated by modulating hmC and mC levels, we performed knock-down of *Tet1* in *Dnmt* TKO cells (Supplementary Fig. 10a). We found that all the tested transcriptional effects by knockdown of *Tet1* were similar in *Dnmt* TKO and normal ES cells (Fig. 3c and Supplementary Fig. 10b), indicating that the effects are independent of catalytic activity. However, we cannot rule out that TET1-dependent modulation of hmC and mC might contribute to transcriptional fine-tuning at some target genes. Taken together, these results indicate that TET1 can contribute to transcriptional repression, and to a minor extent



a

No. of peptides	TET1		TET2	
	Flag	HA	Flag	HA
SIN3A	34	35	0	0
ARID4B	11	10	0	0
HDAC1	4	7	2	2
RBBP7	9	6	0	0
HDAC2	9	4	0	0
SAP130	3	3	0	0
SUDS3	4	2	0	0
SAP30	3	2	0	0

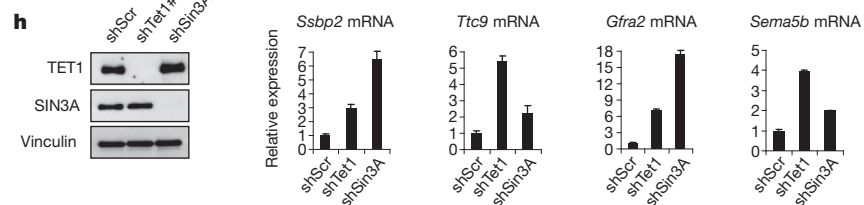
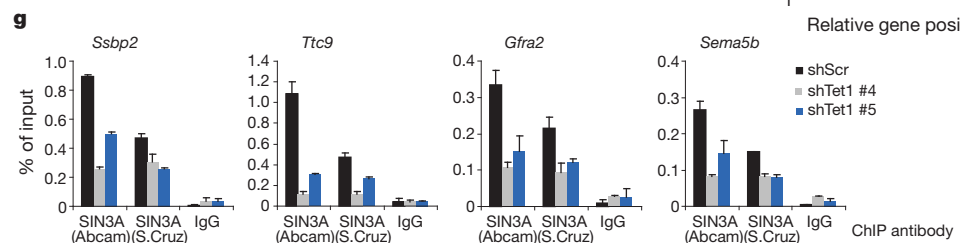
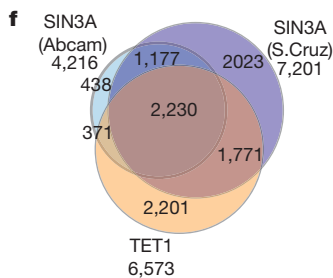
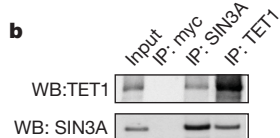
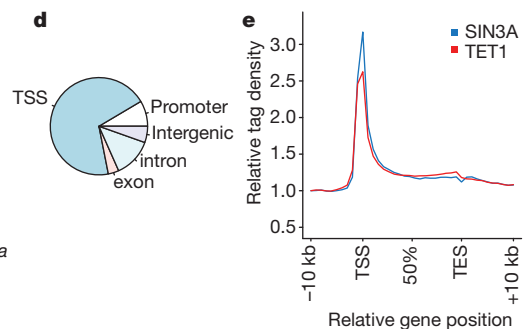
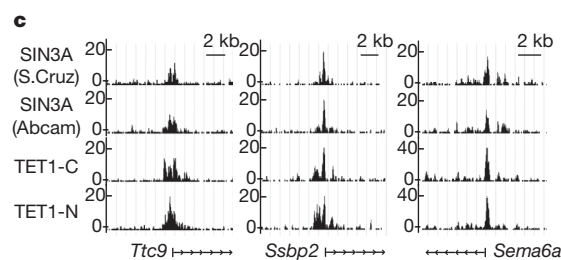


Figure 4 | TET1 interacts with SIN3A. **a**, Peptides identified by mass spectrometry from anti-Flag and tandem anti-Flag-HA purification of Flag-HA-TET1 and Flag-HA-TET2 stably expressed in 293 cells. The presented proteins are all part of the SIN3A complex¹⁷. **b**, Antibodies specific for TET1, SIN3A and c-Myc (negative control) were used for immunoprecipitation (IP) and western blot (WB) using nuclear extracts from mouse ES cells. Input represents 8%. **c**, Examples of SIN3A and TET1 ChIP-seq results in mouse ES cells. **d**, Diagram illustrating the overall distribution of SIN3A binding sites into

Figure 3 | Knockdown of *Tet1* in ES cells affects transcription. **a**, Microarray analyses were performed in control (shScr) and *Tet1* knockdown cells (shTet1#4 and shTet1#5) in triplicates. Venn diagram showing overlap between TET1-bound genes, and genes up- or downregulated by both shRNAs using a cut-off of FDR < 0.05. **b**, qRT-PCR validation of selected genes. **c**, Genes that were found upregulated or downregulated by *Tet1* knockdown show similar regulation in *Dnmt* TKO ES cells. All error bars denote s.d., $n = 3$.

also transcriptional activation, and that the majority of TET1-mediated transcriptional effects are independent of conversion of mC to hmC.

TET1 associates with the SIN3A complex

The mechanism by which TET1 contributes to transcriptional repression is unknown. Although we find an extensive overlap between TET1 and PcG target genes, we have not been able to detect a physical interaction of TET1 with PcG proteins. Therefore, we purified proteins associated with double-epitope Flag-haemagglutinin (Flag-HA)-tagged TET1 expressed in HEK293 cells. This purification led to the identification of SIN3A and several other core components of the SIN3A co-repressor complex, which we did not find associated with the TET2 hydroxylase (Fig. 4a). The SIN3A co-repressor complex is thought to contribute to transcriptional repression by mediating histone deacetylation¹⁷. We validated the interaction between SIN3A and TET1 *in vivo* by co-immunoprecipitation of endogenous proteins with and without the DNA intercalating agent ethidium bromide (Fig. 4b and Supplementary Fig. 11a). Furthermore, TET1 expressed as a fusion protein with the GAL4 DNA binding domain was sufficient

TSS (± 1 kb), promoter (-1 to -5 kb), exon, intron and intergenic regions. **e**, Mean distribution of tags across gene bodies for SIN3A and TET1. **f**, Venn diagram illustrating a significant ($P < 10^{-8}$) overlap between TET1 and SIN3A target genes (FDR < 0.01). **g**, ChIP-qPCR in control or *Tet1* knockdown cells (shTet1#4 and shTet1#5). **h**, Left panel, western blot illustrating knockdown efficiencies of TET1 and SIN3A. Right panel, genes that are upregulated by *Tet1* knockdown are also de-repressed by *Sin3a* knockdown. All error bars denote s.d., $n = 3$.

to recruit SIN3A to the GAL4 DNA binding sites *in vivo* (Supplementary Fig. 11b–e).

To understand if SIN3A also colocalizes with TET1 on target genes, we performed ChIP-seq analysis using two different commercial antibodies to SIN3A (Fig. 4c, Supplementary Table 1). This analysis showed that SIN3A has a similar binding profile as TET1 (Fig. 4d, e and Supplementary Fig. 4), and that TET1 and SIN3A display a significant overlap of target genes (Fig. 4f and Supplementary Fig. 12a). Moreover, ChIP experiments showed that TET1 contributes significantly to the recruitment of SIN3A (Fig. 4g), whereas depletion of SIN3A had no or modest effect on TET1 binding to tested target genes (Supplementary Fig. 12b). To understand if SIN3A is required for the silencing of TET1 repressed genes, we performed gene expression analysis of *Sin3A* knockdown cells (Supplementary Fig. 12c and Supplementary Table 4). Here we found an extensive overlap between genes with increased expression after *Tet1* and *Sin3A* knockdown that are also directly bound by both TET1 and SIN3A (Supplementary Fig. 12d). This implies that SIN3A is required for the repression of a subset of TET1 target genes that show increased expression upon TET1 downregulation (Fig. 4h).

Discussion

One of the major findings presented in this paper is that TET1 localizes to gene bodies and TSSs of a large number of genes and is particularly enriched on genes with high CpG-content. In contrast to the global pattern of mC, which is found predominantly in low CpG density regions, we found that hmC colocalizes with TET1 at high and intermediate CpG-content sequences. This finding indicates that TET1 could have an important role in the metabolism of mC at CpG-rich sequences by converting it to hmC. Statistically significant hmC levels were not detected around the TSS at the majority of TET1 target genes. It is possible that these genes are not methylated and therefore cannot be subsequently hydroxymethylated. Alternatively, it is tempting to speculate that low and stochastically placed methylations on these CpG-rich genes are passively eliminated through replication in rapidly dividing ES cells, following TET1-mediated hydroxylation. If so, the generated hmC will most likely not be detected by DIP-analyses because it will only occur in few cells in the total cell population. In this way the role of TET1 would be to remove aberrant stochastic DNA methylation and contribute to regulating DNA methylation fidelity in ES cells. However, we also found a large number of hmC-positive genes and, interestingly, many of these become hypermethylated in differentiated cells, for example, *Dazl*, *Hormad1*, *Sycp1* and *Sycp2* (ref. 2; Supplementary Table 2 and data not shown). This suggests a dual biological role of TET1, one in which it removes aberrant DNA methylation and another that ensures the timely DNA methylation and silencing of target genes during differentiation.

We also provide evidence that TET1 has a role in transcriptional repression. Interestingly, downregulation of TET1 in *Dnmt* TKO ES cells leads to upregulation of the same genes as observed in wild-type ES cells, indicating that the repressive function of TET1 is independent of its catalytic activity. We found that TET1 interacts with the SIN3A complex and the extensive colocalization of TET1 and the SIN3A co-repressor complex at target genes suggests that SIN3A has an important function in TET1-mediated gene repression.

In summary, our results indicate that TET1 is required for the timely expression of genes during development. We propose that TET1 by converting mC to hmC serves an important function in the regulation of DNA methylation fidelity. In turn this conversion may lead to a reduction of DNA methylation at CpG-rich gene regulatory sequences. Thus, loss of function of the TET proteins would promote the stochastic hypermethylation of promoters leading to deregulation of transcription and differentiation. Interestingly, the related TET2 oxygenase is frequently mutated in a variety of haematopoietic neoplasms supporting an important role of conversion of mC to hmC in cellular homeostasis^{18,19}.

METHODS SUMMARY

Cell culture. Low passage (p17) E14TG2a.4 feeder independent ES cells were grown on 0.1% gelatin-coated plates in standard ES medium²⁰. Recombinant lentiviruses encoding *Tet1* or *Sin3A* shRNA were produced by standard methods. **Generation of antibodies to murine TET1 and hydroxymethylcytosine.** Polyclonal antibodies were generated by immunizing rabbits with affinity-purified bacterially expressed glutathione-S-transferase (GST) fusion proteins GST-Tet1-N (amino acids 1–308), GST-Tet1-C (amino acids 1739–2039) and hydroxymethylcytosine coupled to BSA. The antisera were absorbed on GST or mC-ovalbumin, respectively, and subsequently affinity-purified on the antigens. **ChIP/DIP assays and ChIP/DIP-sequencing.** Chromatin immunoprecipitation assays (ChIP) were performed and analysed as described previously²¹. hme/me-DIP assays were performed as described²². For ChIP-seq analysis, the DNA obtained from the ChIP assays were adaptor-ligated and amplified using a kit from Illumina (IP-102-1001). The amplified DNA from hme/me-DIP or ChIP-seq experiments was analysed by Solexa/Illumina high-throughput sequencing. The tags were mapped to the mouse genome (assembly mm9) with Bowtie or the Solexa Analysis Pipeline. Peak detection and binding analysis were performed using the CisGenome program²³ or MEDIPS²⁴. Chromosomal positions (peaks) were annotated to the RefSeq database.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 October 2010; accepted 28 March 2011.

Published online 13 April 2011.

1. Fouse, S. D. *et al.* Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2**, 160–169 (2008).
2. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
3. Mohn, F. *et al.* Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30**, 755–766 (2008).
4. Saxonov, S., Berg, P. & Brutlag, D. L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl Acad. Sci. USA* **103**, 1412–1417 (2006).
5. Takai, D. & Jones, P. A. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl Acad. Sci. USA* **99**, 3740–3745 (2002).
6. Suzuki, M. M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nature Rev. Genet.* **9**, 465–476 (2008).
7. Cedar, H. & Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Rev. Genet.* **10**, 295–304 (2009).
8. Gal-Yam, E. N., Saito, Y., Egger, G. & Jones, P. A. Cancer epigenetics: modifications, screening, and therapy. *Annu. Rev. Med.* **59**, 267–280 (2008).
9. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
10. Kriaucinis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
11. Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F. & Leonhardt, H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **38**, e181 (2010).
12. Koh, K. P. *et al.* Tet1 and tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
13. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
14. Tsumura, A. *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).
15. Maunakea, A. K. *et al.* Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**, 253–257 (2010).
16. Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotechnol.* **29**, 68–72 (2011).
17. Grzenda, A., Lomber, G., Zhang, J. S. & Urrutia, R. Sin3: master scaffold and transcriptional corepressor. *Biochim. Biophys. Acta* **1789**, 443–450 (2009).
18. Mohamedali, A. M. *et al.* Novel TET2 mutations associated with UPD4q24 in myelodysplastic syndrome. *J. Clin. Oncol.* **27**, 4002–4006 (2009).
19. Delhommeau, F. *et al.* Mutation in TET2 in myeloid cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
20. Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M. & Helin, K. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol. Cell. Biol.* **27**, 3769–3779 (2007).
21. Pasini, D. *et al.* JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature* **464**, 306–310 (2010).
22. Weber, M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genet.* **37**, 853–862 (2005).
23. Ji, H. *et al.* An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nature Biotechnol.* **26**, 1293–1300 (2008).

24. Chavez, L. *et al.* Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. *Genome Res.* **20**, 1441–1450 (2010).
25. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
26. Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
27. Marson, A. *et al.* Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**, 521–533 (2008).
28. Erlanger, B. F. & Beiser, S. M. Antibodies specific for ribonucleosides and ribonucleotides and their reaction with DNA. *Proc. Natl Acad. Sci. USA* **52**, 68–74 (1964).
29. Rhead, B. *et al.* The UCSC Genome Browser database: update 2010. *Nucleic Acids Res.* **38**, D613–D619 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank U. Toftegaard for excellent technical help, M. Okano for the donation of TKO ES cells, and members of the Helin lab for discussions. M.T.P. was supported by a fellowship from the Danish Cancer Society. J.R. is a senior research

fellow of the Wellcome Trust. The work in the Helin lab was supported by grants from the Excellence Program of the University of Copenhagen, the Danish National Research Foundation, the Danish Cancer Society, the Lundbeck foundation, the Novo Nordisk Foundation, and the Danish Medical Research Council.

Author Contributions K.W. performed the major part of experiments in Figs 1, 3, 4a, b, h and Supplementary Figs 1a–c, 2, 3, 5, 7a, 9a–d, 10a, 11 and 12c, d. J.C. developed and characterized the new reagents used in this study, and participated in most experiments. M.T.P. performed the major part of experiments in Figs 2, 4c, g and Supplementary Figs 1d, 6b, 7b, c, 8, 10b and 12a, b. J.V.J. performed bioinformatics analyses. P.A.C.C. assisted in characterizing reagents. J.R. performed the mass spectrometry analysis. J.C. and K.H. supervised the project and all authors contributed to the writing of the manuscript.

Author Information ChIP-seq and gene expression data are available at the Gene Expression Omnibus (GEO) under accession GSE24843. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.H. (kristian.helin@bric.ku.dk).

METHODS

Cell culture. Low passage (p17) E14TG2a.4 feeder independent ES cells were grown on 0.1% gelatin-coated plates in Glasgow medium (Sigma) supplemented with glutamine (Gibco), nonessential amino acids (Gibco), sodium pyruvate (Gibco), 50 μ M β -mercaptoethanol, and 15% fetal bovine serum (HyClone) in the presence of leukaemia inhibitory factor (LIF). Recombinant lentiviruses encoding *Tet1* and *Sin3A* shRNA were produced by standard methods employing co-transfection of pLKO.1 shRNA and packaging vectors in 293FT cells. shRNA-transduced ES cells were selected 36 h post transduction with 2 μ g per ml of puromycin for 72 h. For *Sin3A* knockdown, cells were harvested after 48 h to minimize differentiation. *Tet1* shRNAs had the following sequences, shTet1#3: 5'-tgtagaccatcactgttcgac-3', shTet1#4: 5'-tcactactctcacttagtg-3', shTet1#5: 5'-agagaacctgtgcatcagat-3', shTet1#A: 5'-gcagatggcgtgacacaaat-3' and shTet1#B: 5'-gctcatggagactggttgg-3'. *Sin3A* shRNA had the following sequence, shSin3A#73: 5'-gctgttcgattgctctaaa-3'.

Cloning procedures. The open-reading frames (ORF) of mouse *Tet1* and *Tet2* were amplified by PCR using cDNA from mouse ES cells or LPS-stimulated RAW264.7 mouse macrophages as template, respectively. The amplified fragments were cloned into the pCR8/GW gateway entry vector (Invitrogen), and the DNA sequence was verified by sequencing. Coding errors according to the GenBank reference sequences of mouse *Tet1* and *Tet2* were corrected by site-directed mutagenesis. To generate expression vectors, the appropriate entry clones were transferred into gateway-compatible pCDNA5 TO Flag-HA. shRNA constructs targeting *Tet1* were constructed in pLKO.1. shRNAs targeting murine *Sin3A* were obtained from Sigma-Aldrich.

Generation of antibodies to mouse TET1 and hydroxymethylcytosine. Polyclonal antibodies were generated by immunizing rabbits with affinity-purified bacterially expressed GST-Tet1-N (amino acids 1–308) and GST-Tet1-C (amino acids 1739–2039). The antibodies were absorbed on GST-coupled cyanogen bromide-activated Sepharose (GE Healthcare) and subsequently affinity purified using Sepharose coupled with GST-Tet1-N or GST-Tet1-C. Antibody specificity was confirmed by immunoblotting and immunoprecipitation. To generate antibodies against hydroxymethylcytosine, 5-hydroxymethylcytidine (Berry & Associates), was covalently coupled to BSA essentially as described²⁸ and used for immunization of rabbits. Affinity-purified anti-hydroxymethylcytosine (hmC) antibodies were produced by column absorption of the rabbit antisera on methylcytidine-ovalbumin coupled to cyanogen bromide-activated Sepharose followed by column-affinity purification on hydroxymethylcytidine-ovalbumin coupled to Sepharose. The antibodies were eluted with 0.1 M glycine-HCl, neutralised, dialysed against PBS and stored at -80°C . The specificity of the purified anti-hmC antibodies were analysed by ELISA and in hme-DIP assays. For the hme-DIP assays, synthetic 300-base-pair probes incorporating 5, 20 and 100% hmC or mC, respectively, were amplified by PCR using pCR8/GW (nucleotides 701–1000) as template. The probes (0.001 ng) were spiked into the hmeC/meC reactions containing 1 μ g of sonicated ES DNA. Antibody reactivity with the probes was detected by qPCR.

Purification of TET1 and TET2 complexes. To isolate TET1 and TET2-containing complexes, two-step affinity purification was performed followed by mass spectrometry analysis. Nuclear extracts (250–500 mg, 3×10^9 cells) from Flp-In-T-REX-293 cell lines expressing Flag-HA-tagged murine TET1 or TET2 were precleared and incubated with a 700 μ l packed volume of anti-Flag beads (anti-Flag M2-agarose, Sigma) overnight at 4°C with rotation. The beads were collected by centrifugation at 700g for 5 min and washed six times with 40 \times resin bed volume of buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1.5 mM MgCl_2 , 0.2 mM EDTA, 10% glycerol, 0.2 mM PMSF, 1 mM DTT, 1 μ g ml $^{-1}$ aprotinin and 1 μ g ml $^{-1}$ leupeptin). The beads were transferred into a 10-ml poly-prep chromatography column (Bio-Rad) and complexes were then eluted five times after 10 min of incubation using one resin bed volume of buffer A supplemented with 0.5 μ g μ l $^{-1}$ Flag peptide. The eluate was subjected to a second round of purification using an antibody against the HA-tag. The Flag-IP elute was incubated with 200 μ l of a 50% slurry of HA-beads overnight. The beads were washed four times with buffer A and eluted with 100 μ l buffer A supplemented with 1 μ g μ l $^{-1}$ HA peptide for 2 h. The samples were boiled in SDS loading buffer and run shortly into a SDS-PAGE gel in order to remove the Flag and HA peptide and other contaminations. A gel slice containing the purified proteins was isolated for mass spectrometry analysis.

ChIP/DIP assays and ChIP/DIP-seq. Chromatin immunoprecipitation assays (ChIP) were performed and analysed as previously described²¹. The antibodies used were anti-mSin3A (Abcam AB3479, Santa Cruz sc-994X) and the antibodies to TET1 described above. ES cell DNA was sonicated to an average size between 300 and 600 bp. Adaptor-ligated libraries for hmC or mC DNA immunoprecipitation assays (hm-DIP/me-DIP) were constructed using the NEBNext DNA Sample Prep Master Mix, NEB combined with Illumina adaptors. hme/me-DIP

assays were performed as described²² using 1 μ g of denatured sonicated or adaptor-ligated DNA in 100 μ l of binding buffer and 0.1–4 μ g of affinity-purified rabbit hmC antibody or monoclonal mC antibody (Eurogentec BI-MECY-0500). The samples were incubated for four hours at 4°C before addition of 10 μ l of anti-rabbit/mouse Dynabeads (Invitrogen). After 2 h of incubation, the samples were washed four times and bound DNA was eluted by incubation for one hour at 55°C in 100 μ l of 50 mM Tris-HCl, 10 mM EDTA, 0.5% SDS and 20 μ g proteinase K. The DNA was purified using a QIAquick PCR purification kit (Qiagen) and amplified by 16 cycles of PCR. For the MeCAP (methylated DNA capture by affinity purification) experiments, the MethylCap kit (Diagenode) was used according to manufacturer's instructions. For ChIP-seq analysis, the DNA obtained from the ChIP assays were adaptor-ligated and amplified using a kit from Illumina (IP-102-1001). The amplified DNA from hme/me-DIP or ChIP-seq experiments was analysed by Solexa/Illumina high-throughput sequencing. After prefiltering the raw data by removing sequenced adapters and low quality reads, the tags were mapped to the mouse genome (assembly mm9) with the Bowtie alignment tool. To avoid any PCR bias we allowed only one read per chromosomal position (unless otherwise specified) thus eliminating spurious spikes. Peak detection were performed in the CisGenome program²³ at an FDR cut-off value <0.1 or <0.01 as indicated in the text. IgG was used as control for normalization. Venn diagram analysis was performed with Galaxy browser (www.galaxy.psu.edu). Most standard peak detection programs are typically optimized for transcription binding site data and anticipate a defined narrow bell-shaped density profile. However, for epigenetics data, such as mC and hmC, the peaks tends to be broad and low-intensity, thus requiring a different peak detection program. We used the MEDIPS tool²⁴ (bin size = 50, fragment length = 250, frame size = 500, step = 250) to detect significant enrichment of signal (reads per million, rpm) relative to a control (*Dnmt* TKO DIP) and an input (IgG DIP) sample at an FDR cut-off value <0.1 and a minimal enrichment of ratio >5 . For the MEDIPS analysis the reads were not limited to one read per chromosomal position, and the total length of the mapped reads were extended in the direction of the 3'-end to a total length of 250 bases, which was our estimate of the mean fragment length. Chromosomal positions (peaks) were annotated to the RefSeq database (mm9) using the UCSC "refFlat" table²⁹. Genes not uniquely mapped to the genome were excluded. Signal vs CpG plot: for the signal vs CpG plots the MEDIPS calculated rpm and CpG (CpG values from transformed "coupling" factors) values were used. To avoid redundancy only the longest transcript variant of each gene was used to define chromosomal locations of promoter, TSS, exons and introns. For each bin (non-overlapping) MEDIPS determines the number of overlapping reads and the CpG content. For a specified region of interest (ROI), for example, an exon, the mean rpm and CpG content of the bins within the range was calculated. The distribution of CpG content within the different genomic categories are distinctly different for example, with the TSS region showing the known bimodal distribution. To depict the rpm as a function of CpG-content, the mean rpm values were stratified according to CpG-content (1% resolution) and the mean of the mean rpms within each stratus calculated. Due to variability in the size of ROIs (except for the genome-wide analysis), the plots for the different genomic categories are not directly comparable. Wiggle-based plots: to avoid redundancy, the longest transcript variant of each gene in the RefSeq database was used as reference. In total the chromosomal mappings of 21,513 unique genes were used. The filtered alignment files were converted to bigWig files from which the tag count information was extracted using unix tools from the UCSC website. Gene Body plots: 40 non-overlapping windows with average tag number per base were calculated for each gene. 10 kb upstream of TSS and 10 kb downstream of transcription end site (TES) was divided into of windows of size 1 kb. Between TSS and TES each gene were divided into 20 windows of equal (gene-specific) size and the average counts was calculated. All statistics and plotting were done using the statistical program R.

mRNA expression analysis. For expression analysis, total RNA was purified from murine embryonic stem cells using RNeasy (Qiagen). The RNA was reverse transcribed using TaqMan reverse transcription reagents from ABI, according to the manufacturer's instructions. For RNA quantification, reversed-transcribed total RNA was analysed by real-time PCR using SYBR Green PCR Master Mix (Fermentas) and an ABI prism 7700 Sequence Detection system. All reactions were analysed in triplicates. Primer sequences are listed in Supplementary Figure 13 and Supplementary Fig. 14. For microarray analysis, RNA was extracted with the RNeasy Plus RNA extraction kit (Qiagen). RNA was hybridized on mouse Gene 1.0 ST arrays by the RH Microarray Center at Rigshospitalet, Copenhagen, following Affymetrix procedures and analysis. Gene expression analyses of RNA from shScr, shTet1#4, shTet1#5 and shSin3A#73 cells were performed in triplicates and in the subsequent data analysis FDR values <0.05 was used.

Unbound or distant planetary mass population detected by gravitational microlensing

The Microlensing Observations in Astrophysics (MOA) Collaboration & The Optical Gravitational Lensing Experiment (OGLE) Collaboration*

Since 1995, more than 500 exoplanets have been detected using different techniques^{1,2}, of which 12 were detected with gravitational microlensing^{3,4}. Most of these are gravitationally bound to their host stars. There is some evidence of free-floating planetary-mass objects in young star-forming regions^{5–8}, but these objects are limited to massive objects of 3 to 15 Jupiter masses with large uncertainties in photometric mass estimates and their abundance. Here, we report the discovery of a population of unbound or distant Jupiter-mass objects, which are almost twice ($1.8^{+1.7}_{-0.8}$) as common as main-sequence stars, based on two years of gravitational microlensing survey observations towards the Galactic Bulge. These planetary-mass objects have no host stars that can be detected within about ten astronomical units by gravitational microlensing. However, a comparison with constraints from direct imaging⁹ suggests that most of these planetary-mass objects are not bound to any host star. An abrupt change in the mass function at about one Jupiter mass favours the idea that their formation process is different from that of stars and brown dwarfs. They may have formed in proto-planetary disks and subsequently scattered into unbound or very distant orbits.

In a gravitational microlensing event, a foreground lens object is detected as a result of the characteristic magnification of a background source star as it passes behind the gravitational field of the lens¹⁰. The lens object is detected by means of its mass and not its luminosity. The duration of the magnification is parameterized by the Einstein radius crossing time, $t_E \approx \sqrt{M/M_J}$ days, where $M_J = 9.5 \times 10^{-4} M_\odot$ is Jupiter's mass (M_\odot is the mass of the Sun). Thus, microlensing can detect faint planetary mass objects—which are either unbound to any host star^{11,12} or are in very wide orbits¹³—as short-timescale events with $t_E < 2$ days. Although t_E also depends on the distance and transverse velocity of the lens (see Supplementary Information), the observed t_E distribution can be used as a statistical probe of the mass function of the lens objects because the spatial and velocity distributions in the Galactic disk and bulge are reasonably well known.

The Microlensing Observations in Astrophysics (MOA)¹⁴ and Optical Gravitational Lensing Experiment (OGLE)¹⁵ groups both conduct microlensing surveys towards the Galactic Bulge. The second phase of MOA, called MOA-II, carries out a high-cadence photometric survey of 50 million stars in bulge fields with a cadence of 10–50 min. This strategy enables MOA to detect very short events with $t_E < 2$ days, which were quite rare in previous microlensing surveys that had lower cadences^{12,14,16,17}.

In this analysis of the 2006–2007 MOA-II data set, light curves of genuine microlensing events were distinguished from intrinsic variables and artefacts by several empirical criteria, which have been developed in previous microlensing surveys^{14–19}. The light curves must have a single brightening episode of more than three consecutive significant data points with a constant baseline, and should be well fitted by a theoretical microlensing model⁹ with a well constrained t_E (see the Supplementary Information).

Although there are a thousand microlensing events in this sample, only 474 well characterized events passed our strict selection criteria. Ten of these events have $t_E < 2$ days (see Fig. 1 and Table 1)—thus indicating planetary-mass lenses. We have confirmed that this event sample has no significant contamination by possible background effects including: (1) cosmic-ray hits, (2) fast-moving objects, (3) cataclysmic variables, (4) background supernovae, (5) binary microlensing events and (6) microlensing by high-velocity stars and Galactic-halo stellar remnants. For example, effect (1) is excluded because cosmic rays never hit the same place in four consecutive images, microlensing model fits for effects (2) to (5) produce a high χ^2 and unphysical values of parameters, and effect (6) is excluded by proper-motion and radial-velocity observations (see Supplementary Information). After the MOA event selection was complete, the MOA group requested additional independent light-curve data of these short events from the OGLE group. Seven of the ten events with $t_E < 2$ days were also observed by OGLE-III¹⁵, and none of them have any other brightening in the eight-year OGLE-III light curves. For six of these seven events, there are OGLE data obtained during the lensing event that confirm the predictions of the MOA microlensing models. Thus, the OGLE data confirms the microlensing interpretation of these short events.

The detection efficiencies for this analysis were estimated with a Monte Carlo simulation^{14,17}. We simulated 20 million artificial events to evaluate the detection efficiency as a function of t_E , yielding $\varepsilon(t_E) \approx 1\%$, 3% , 5% , 10% , 15% and 10% at $t_E = 0.3$, 1 , 2 , 10 , 30 and 100 days, respectively. The details of the efficiency calculations and consistency tests of the selected event distribution are discussed in the Supplementary Information.

The observed t_E distribution is compared to two mass function models in Fig. 2. A model t_E distribution, $\Phi(t_E)$, can be calculated for an assumed mass function with a standard Galactic mass density and velocity model^{14,17,20}. We consider two mass functions. The first is a broken power law^{20,21} $dN/dM = M^{-\alpha}$, with power indices of $\alpha_1 = 2.0$ for $0.7 \leq M/M_\odot \leq 1$, $\alpha_2 = 1.3$ for $0.08 \leq M/M_\odot \leq 0.7$ and α_3 as a fitting parameter for the brown dwarf regime $0.01 \leq M/M_\odot \leq 0.08$. The second is a log-normal function²² $dN/d\log M = \exp[(\log M - \log M_c)^2 / (2\sigma^2)]$ with a mean mass M_c and a width in $\log M$ of σ , for $0 \leq M/M_\odot \leq 1.0$. For both mass functions, we assume that stars that were initially above $1M_\odot$ have evolved into stellar remnants—white dwarfs, neutron stars or black holes, depending on their initial masses²³ (see Fig. 2 and Supplementary Table 3).

The mass functions were constrained by a likelihood analysis, with the likelihood function given by the product of the model probability $\Phi(t_E)$ of finding $N_{\text{obs}} = 474$ events with each of the observed $t_{E,i}$ that is: $L = \prod_i^{N_{\text{obs}}} \Phi(t_{E,i}) \varepsilon(t_{E,i})$.

We evaluated the likelihood distributions for these mass functions both with and without the $t_E < 2$ events, but the inclusion of the events with $t_E < 2$ days makes little difference. The results are shown in Supplementary Table 3 and Supplementary Figs 6 and 7. Figure 2 indicates that both models match the data well for $t_E \geq 2$ days, but at

*Lists of participants and their affiliations appear at the end of the paper.

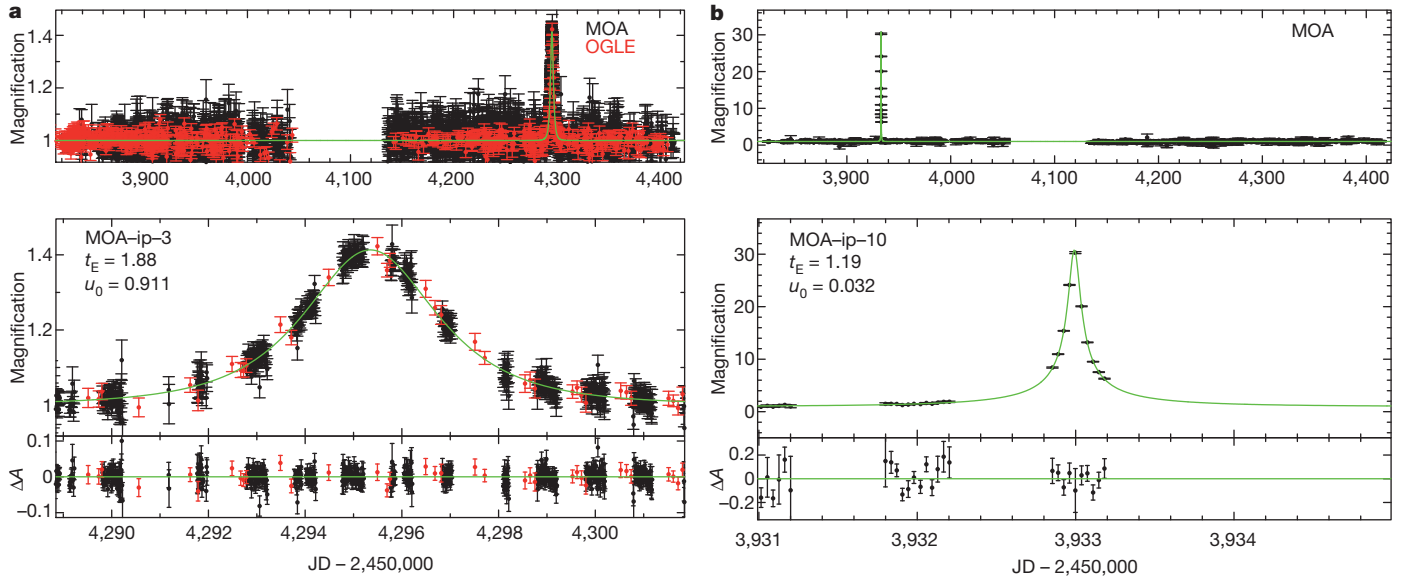


Figure 1 | Light curves of event MOA-ip-3 and event MOA-ip-10. These have the highest signal-to-noise ratio of the ten microlensing events with $t_E < 2$ days (see Supplementary Fig. 1 for the others). MOA data are in black and OGLE data are in red, with error bars indicating the s.e.m. **a**, MOA-ip-3 light curve; **b**, MOA-ip-10 light curve. The green lines represent the best-fit microlensing model light curves. For each event, the upper panel shows the full two-year light curve, the middle panel is a close-up of the light-curve peak, and the bottom panel shows the residuals from the best-fit model in units of the magnification, ΔA . u_0 indicates the source-lens impact parameter in units of the Einstein radius. The second phase of MOA, MOA-II, carried out a very-high-cadence photometric survey of 50 million stars in 22 bulge fields (of

2.2 deg² each) with a 1.8-m telescope at Mt John Observatory in New Zealand. MOA detects 500–600 microlensing events during eight months observation every year. In 2006–2007, MOA observed two central bulge fields every 10 min, and other bulge fields with a 50 min cadence, which resulted in about 8,250 and 1,660–2,980 images, respectively. This strategy enabled MOA to detect very short events with $t_E < 2$ days. Since 2002, the OGLE-III survey has monitored the bulge with the 1.3-m Warsaw telescope at Las Campanas Observatory, Chile, with a smaller field-of-view but better astronomical seeing than MOA. The OGLE-III observing cadence was 1–2 observations per night, but the OGLE photometry is usually more precise and fills gaps in the MOA light curves owing to the difference in longitude.

$t_E < 2$ days, the ten observed events are well above the model predictions. The power-law and log-normal models predict 1.5 and 2.5 events with $t_E < 2$ days, respectively, and the corresponding Poisson probabilities for the ten observed events are 4×10^{-6} and 3×10^{-4} . Thus, we feel confident in adding a new planetary-mass population.

For simplicity, we chose a planetary-mass function model with a δ -function in mass M_{PL} and a fraction of all objects in the planetary-mass population Φ_{PL} . The values of $(M_{PL}/M_\odot, \Phi_{PL})$ derived from the likelihood analysis are $(1.1^{+1.2}_{-0.6} \times 10^{-3}, 0.49^{+0.13}_{-0.13})$ and $(0.83^{+0.96}_{-0.51} \times 10^{-3}, 0.46^{+0.17}_{-0.15})$ for the power-law and log-normal models, respectively. The contours are shown in Fig. 3. Both models for $\Phi(t_E)$ provide good fits to the entire t_E distribution, as shown in Fig. 2. The power-law and log-normal models imply $1.9^{+1.3}_{-0.8}$ and $1.8^{+0.7}_{-0.8}$ times as many unbound or distant Jupiter-mass objects as main-sequence stars in the mass range $0.08 < M/M_\odot < 1.0$, respectively. These planetary mass objects are at least 1.5 times as frequent as planets with host stars (see Supplementary Information). We tested a third mass function that has fewer massive

stars and brown dwarfs²⁴, and found that the resultant planetary-mass function parameters are consistent with the above values (see the Supplementary Information).

The lenses for these short events could be either free-floating planets or planets with wide separations of more than about ten astronomical units (AU) from their host stars, for which we cannot detect the host star in the light curves²⁵. However, direct imaging, with adaptive optics, of planets orbiting young stars places upper limits on planets at wide separations. The Gemini Planet Imager has set upper limits⁹ on the number of stars with Jupiter-mass planets at semi-major axes of 10–500 AU. From these results, we estimate that < 0.4 of the 1.8 planetary-mass objects per star are likely to be bound to stars at orbital separations of < 500 AU (see Supplementary Information section 8). Hence, more than 75% of these planetary mass objects are probably unbound to stars if their typical mass is a Jupiter-mass or more.

Because the δ -function planetary models are not likely to be realistic, we also tested a fourth mass function that is identical to the first,

Table 1 | Microlensing parameters of short-timescale events

ID	Field	Right ascension, α (2000)	Declination, δ (2000)	N_{t_E}	t_0 (JD')	t_E (days)	u_0 (R_E)	A_{max}	I_s (mag)	d_{min} (R_{E*})
MOA-ip-1	gb1-4	17 h 46 min 24.506 s	−34° 30' 36.82''	9	3883.24171	0.73 ± 0.08	0.028 ± 0.003	35.6	19.7	7.0
MOA-ip-2	gb4-3	17 h 52 min 34.143 s	−30° 54' 14.25''	28	4223.88851	0.49 ± 0.10	0.400 ± 0.212	2.6	17.9	3.3
MOA-ip-3	gb5-7	17 h 54 min 58.325 s	−29° 38' 20.68''	170	4295.34720	1.88 ± 0.12	0.911 ± 0.096	1.4	17.2	3.6
MOA-ip-4	gb5-8	17 h 54 min 24.543 s	−29° 13' 29.39''	81	3961.38803	1.48 ± 0.12	0.271 ± 0.061	3.8	19.2	3.1
MOA-ip-5	gb9-2	17 h 57 min 17.008 s	−29° 02' 33.59''	69	4169.60907	1.62 ± 0.69	0.126 ± 0.159	8.0	19.2	2.4
MOA-ip-6	gb9-4	17 h 59 min 19.977 s	−29° 31' 24.70''	27	4189.49214	1.78 ± 0.24	0.499 ± 0.122	2.2	18.3	4.8
MOA-ip-7	gb9-5	17 h 57 min 36.678 s	−29° 59' 40.52''	51	4370.69496	1.82 ± 0.87	0.143 ± 0.125	7.0	19.4	5.2
MOA-ip-8	gb9-5	17 h 59 min 34.877 s	−30° 04' 24.04''	47	4013.14052	1.36 ± 0.15	0.103 ± 0.016	9.8	18.8	4.8
MOA-ip-9	gb10-5	17 h 57 min 52.952 s	−28° 16' 56.66''	16	3910.81772	0.96 ± 0.21	0.163 ± 0.058	6.2	19.5	3.4
MOA-ip-10	gb11-9	18 h 09 min 00.076 s	−32° 18' 39.91''	21	3932.99205	1.19 ± 0.04	0.032 ± 0.001	30.8	18.8	15.0

N_{t_E} indicates the number of data points within $t_0 \pm t_E$, and t_0 , t_E , A_{max} and I_s indicate the time of peak magnification, the Einstein radius crossing time, the maximum magnification, and the source star magnitude of the best fit models of the MOA data, respectively. $JD' = JD - 2,450,000$. u_0 and d_{min} indicate the source-lens impact parameter and minimum host star separation in units of the Einstein radii of the planetary mass lens, R_E , and possible host star, R_{E*} , respectively. The errors in t_E and u_0 represent 1 σ limits, d_{min} indicates 2 σ limits. MOA-ip-2, MOA-ip-3 and MOA-ip-10 were alerted as MOA-2007-BLG-144, MOA-2007-BLG-309 and MOA-2006-BLG-098 by the MOA real-time alert system (<http://www.massey.ac.nz/~iabond/moa>).

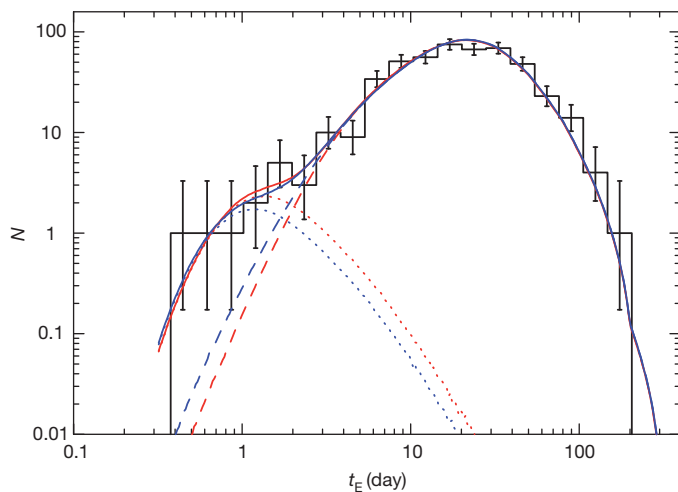


Figure 2 | Observed and theoretical distributions of the event timescale, t_E . The black histogram represents the number N of observed 474 microlensing events in each bin with error bars indicating the s.e.m. The red and blue lines indicate the best-fit models with the power-law and log-normal mass functions, respectively. For both mass functions, we assume that stars that were initially above $1M_\odot$ have evolved into stellar remnants—white dwarfs, neutron stars or black holes, depending on their initial masses. The number of remnants is determined by extending the upper main-sequence power law $\alpha_1 = 2.0$ to $100M_\odot$, and the final remnant mass distributions are given by Gaussians²³ (see Supplementary Table 3). Each model is multiplied by the detection efficiencies. In each model, dashed lines indicate models for stellar, stellar remnant and brown dwarf populations, and the dotted lines represent the planetary-mass population. Solid lines are the sums of these populations, and both models fit the data well.

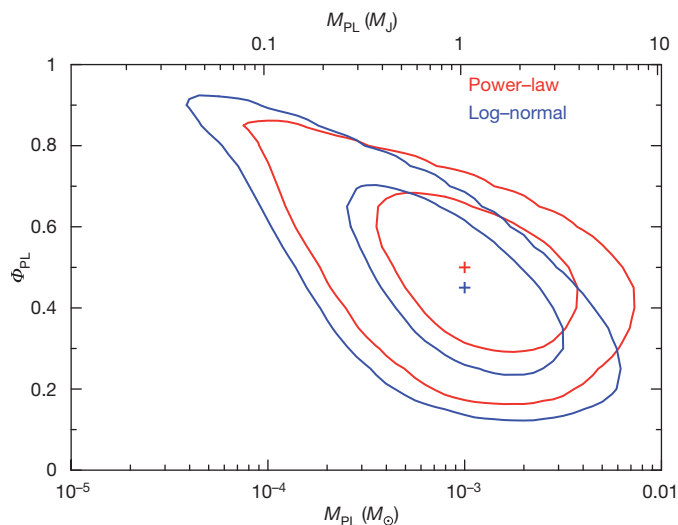


Figure 3 | Likelihood contours for the planetary-mass function parameters. Φ_{PL} indicates the fraction of all objects in the planetary-mass population, not including the brown dwarfs that have planetary mass in the tail of the log-normal mass function. M_{PL} represents their masses. The two sets of contours indicate the 68% and 95% confidence levels. The red and blue curves indicate the power-law and log-normal mass functions, respectively, and crosses indicate the maximum-likelihood points. The top-axis scale is in Jupiter masses and the bottom-axis scale is in solar masses. For the power-law model, the likelihoods are evaluated in the $(\alpha_3, M_{PL}, \Phi_{PL})$ space and projected into the (M_{PL}, Φ_{PL}) plane. The $M_c = 0.12$ and $\sigma_c = 0.76$ parameters are fixed for the log-normal model. The median and 68% confidence intervals of $(M_{PL}/M_\odot, \Phi_{PL})$ are $(1.1^{+1.2}_{-0.6} \times 10^{-3}, 0.49^{+0.13}_{-0.13})$ and $(0.83^{+0.96}_{-0.51} \times 10^{-3}, 0.46^{+0.17}_{-0.15})$ for the power-law and log-normal models, respectively. The results for two models are consistent with each other. The power-law and log-normal models imply $1.9^{+1.3}_{-0.8}$ and $1.8^{+1.7}_{-0.8}$ times as many unbound or distant Jupiter-mass objects as the main-sequence stars. α_3 is consistent with the values derived without planetary population, indicating that brown dwarfs are 0.7 ± 0.3 times as common as main-sequence stars. The numerical values of the models are summarized in Supplementary Table 3.

broken-power-law model except for having a power-law form in the planetary mass regime $M < 0.01M_\odot$. This yields a planetary-mass index of $\alpha_{PL} = 1.3^{+0.3}_{-0.4}$, which is much steeper than the brown dwarf slope of $\alpha_3 = 0.49^{+0.24}_{-0.27}$, indicating that they are distinct populations (see Supplementary Information).

Planet-formation theories predict that dynamical instabilities in planetary systems with multiple giant planets could scatter many of these planets into unbound orbits²⁶, as well as some into large separations²⁷. Recent observations also indicate that planet–planet scattering plays an important part in moving giant planets into short-period orbits^{28,29}. The planetary-mass population that we have identified here may have formed in protoplanetary disks at much smaller separations and then been scattered into unbound or very distant orbits.

Received 19 November 2010; accepted 1 April 2011.

1. Mayor, M. & Queloz, D. A. Jupiter-mass companion to a solar-type star. *Nature* **378**, 355–359 (1995).
2. Schneider, J. *The Extrasolar Planets Encyclopaedia* (<http://exoplanet.eu/>) (CNRS/LUTH—Paris Observatory, 2011).
3. Gaudi, B. S. in *Exoplanets* (ed. Seager, S.) 79–110 (University of Arizona Press, 2011).
4. Sumi, T. *et al.* A cold Neptune-mass planet OGLE-2007-BLG-368Lb: cold Neptunes are common. *Astrophys. J.* **710**, 1641–1653 (2010).
5. Zapatero Osorio, M. R. *et al.* Discovery of young, isolated planetary mass objects in the σ Orionis star cluster. *Science* **290**, 103–107 (2000).
6. Burgess, A. S. M. *et al.* Young T-dwarf candidates in IC 348. *Astron. Astrophys.* **508**, 823–831 (2009).
7. Quanz, S. P. *et al.* Search for very low-mass brown dwarfs and free-floating planetary-mass objects in Taurus. *Astrophys. J.* **708**, 770–784 (2010).
8. Marsh, K. A. *et al.* Deep near-infrared imaging of the ρ Oph cloud core: clues to the origin of the lowest-mass brown dwarfs. *Astrophys. J.* **719**, 550–560 (2010).
9. Lafrenière, D. *et al.* The Gemini deep planet survey. *Astrophys. J.* **670**, 1367–1390 (2007).
10. Paczyński, B. Gravitational microlensing by the galactic halo. *Astrophys. J.* **304**, 1–5 (1986).
11. Liebes, S. Gravitational lenses. *Phys. Rev.* **133**, B835–B844 (1964).
12. Bennett, D. P. *et al.* in *Planets Beyond the Solar System and the Next Generation of Space Missions* (ed. Soderblom, D. R.) ASP Conf. Proc. 119, 95 (Astronomical Society of the Pacific, 1997).
13. Di Stefano, R. & Scalzo, A. A new channel for the detection of planetary systems through microlensing. I. Isolated events due to planet lenses. *Astrophys. J.* **512**, 564–578 (1999).
14. Sumi, T. *et al.* Microlensing optical depth toward the Galactic Bulge from microlensing observations in astrophysics group observations during 2000 with difference image analysis. *Astrophys. J.* **591**, 204–227 (2003).
15. Udalski, A. The Optical Gravitational Lensing Experiment. Real time data analysis systems in the OGLE-III survey. *Acta Astron.* **53**, 291–305 (2003).
16. Udalski, A. *et al.* The Optical Gravitational Lensing Experiment—the optical depth to gravitational microlensing in the direction of the Galactic Bulge. *Acta Astron.* **44**, 165–189 (1994).
17. Alcock, C. *et al.* The MACHO project: microlensing optical depth toward the Galactic Bulge from difference image analysis. *Astrophys. J.* **541**, 734–766 (2000).
18. Hamadache, C. *et al.* Galactic Bulge microlensing optical depth from EROS-2. *Astron. Astrophys.* **454**, 185–199 (2006).
19. Sumi, T. *et al.* Microlensing optical depth toward the Galactic Bulge using bright sources from OGLE-II. *Astrophys. J.* **636**, 240–260 (2006).
20. Han, C. & Gould, A. Stellar contribution to the Galactic Bulge microlensing optical depth. *Astrophys. J.* **592**, 172–175 (2003).
21. Zoccali, M. *et al.* The initial mass function of the Galactic Bulge down to $\sim 0.15M_\odot$. *Astrophys. J.* **530**, 418–428 (2000).
22. Chabrier, G. Galactic stellar and substellar initial mass function. *Publ. Astron. Soc. Pacif.* **115**, 763–795 (2003).
23. Gould, A. Measuring the remnant mass function of the Galactic Bulge. *Astrophys. J.* **535**, 928 (2000).
24. Thies, I. & Kroupa, P. A discontinuity in the low-mass initial mass function. *Astrophys. J.* **671**, 767–780 (2007).
25. Han, C. & Kang, Y. W. Probing the spatial distribution of extrasolar planets with gravitational microlensing. *Astrophys. J.* **596**, 1320–1326 (2003).
26. Veras, D., Crepp, J. R. & Ford, E. B. Formation, survival, and detectability of planets beyond 100 AU. *Astrophys. J.* **696**, 1600–1611 (2009).
27. Marois, C. *et al.* Direct imaging of multiple planets orbiting the star HR 8799. *Science* **322**, 1348–1352 (2008).
28. Winn, J. N. *et al.* Hot stars with hot Jupiters have high obliquities. *Astrophys. J.* **718**, L145 (2010).
29. Howard, A. W. *et al.* The occurrence and mass distribution of close-in super-Earths, Neptunes, and Jupiters. *Science* **330**, 653–655 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The MOA collaboration thanks the JSPS and MEXT of Japan, and the Marsden Fund of New Zealand, for support. D.P.B. acknowledges support by the NSF and NASA. The OGLE collaboration is grateful for funding from the European Research Council Advanced Grants Program.

Author Contributions T.S. and K.K. conducted data reduction and statistical analysis. A.U. produced OGLE-III light curves. I.A.B. generated the extended MOA-II light curves. D.P.B. conducted the detailed analysis of binary events. T.S. wrote the manuscript. D.P.B. and I.A.B. edited the manuscript. All other authors contributed to the observation and maintenance of observational facilities, discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.S. (sumi@ess.sci.osaka-u.ac.jp).

The Microlensing Observations in Astrophysics (MOA) Collaboration

T. Sumi^{1,2}, K. Kamiya², D. P. Bennett³, I. A. Bond⁴, F. Abe², C. S. Botzler⁵, A. Fukui², K. Furusawa², J. B. Hearnshaw⁶, Y. Itow², P. M. Kilmartin⁷, A. Korpela⁸, W. Lin⁴, C. H. Ling⁴, K. Masuda², Y. Matsubara², N. Miyake², M. Motomura², Y. Muraki⁹, M. Nagaya², S. Nakamura², K. Ohnishi¹⁰, T. Okumura², Y. C. Perrott¹¹, N. Rattenbury⁵, To. Saito¹², T. Sako², D. J. Sullivan⁸, W. L. Sweatman⁴, P. J. Tristram⁵ & P. C. M. Yock⁵

The Optical Gravitational Lensing Experiment (OGLE) Collaboration

A. Udalski¹³, M. K. Szymański¹³, M. Kubiak¹³, G. Pietrzyński^{13,14}, R. Poleski¹³, I. Soszyński¹³, Ł. Wyrzykowski¹⁵ & K. Ulaczyk¹³

¹Department of Earth and Space Science, Osaka University, Osaka 560-0043, Japan.

²Solar-Terrestrial Environment Laboratory, Nagoya University, Nagoya 464-8601, Japan.

³Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA.

⁴Institute of Information and Mathematical Sciences, Massey University, Private Bag 102-904, North Shore Mail Centre, Auckland 0745, New Zealand.

⁵Department of Physics, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.

⁶Department of Physics and Astronomy, University of Canterbury, Christchurch 8140, New Zealand.

⁷Mt John University Observatory, University of Canterbury, PO Box 56, Lake Tekapo 8770, New Zealand.

⁸School of Chemical and Physical Sciences, Victoria University, Wellington 6140, New Zealand.

⁹Department of Physics, Konan University, Nishiokamoto 8-9-1, Kobe 658-8501, Japan.

¹⁰Nagano National College of Technology, Nagano 381-8550, Japan.

¹¹Cavendish Laboratory, Cambridge University, J. J. Thomson Avenue, CB3 0HE Cambridge, UK.

¹²Tokyo Metropolitan College of Industrial Technology, Tokyo 116-8523, Japan.

¹³Warsaw University Observatory, Al. Ujazdowskie 4, 00-478 Warszawa, Poland.

¹⁴Universidad de Concepción, Departamento de Astronomía, Casilla 160-C, Concepción, Chile.

¹⁵Institute of Astronomy, Cambridge University, Madingley Road, CB3 0HA Cambridge, UK.

Observation of the antimatter helium-4 nucleus

The STAR Collaboration*

High-energy nuclear collisions create an energy density similar to that of the Universe microseconds after the Big Bang¹; in both cases, matter and antimatter are formed with comparable abundance. However, the relatively short-lived expansion in nuclear collisions allows antimatter to decouple quickly from matter, and avoid annihilation. Thus, a high-energy accelerator of heavy nuclei provides an efficient means of producing and studying antimatter. The antimatter helium-4 nucleus ($^4\bar{\text{He}}$), also known as the anti- α ($\bar{\alpha}$), consists of two antiprotons and two antineutrons (baryon number $B = -4$). It has not been observed previously, although the α -particle was identified a century ago by Rutherford and is present in cosmic radiation at the ten per cent level². Antimatter nuclei with $B < -1$ have been observed only as rare products of interactions at particle accelerators, where the rate of antinucleus production in high-energy collisions decreases by a factor of about 1,000 with each additional antinucleon^{3–5}. Here we report the observation of $^4\bar{\text{He}}$, the heaviest observed antinucleus to date. In total, 18 $^4\bar{\text{He}}$ counts were detected at the STAR experiment at the Relativistic Heavy Ion Collider (RHIC; ref. 6) in 10^9 recorded gold-on-gold (Au+Au) collisions at centre-of-mass energies of 200 GeV and 62 GeV per nucleon–nucleon pair. The yield is consistent with expectations from thermodynamic⁷ and coalescent nucleosynthesis⁸ models, providing an indication of the production rate of even heavier antimatter nuclei and a benchmark for possible future observations of $^4\bar{\text{He}}$ in cosmic radiation.

In 1928, the existence of negative energy states of electrons was predicted⁹ on the basis of the application of symmetry principles to quantum mechanics, but these states were only recognised to be antimatter after the discovery¹⁰ of the positron (the antielectron) in cosmic radiation four years later. The predicted antiprotons¹¹ and antineutrons¹² were observed in 1955, followed by antideuterons (\bar{d}), antitritons ($^3\bar{\text{H}}$), and antihelium-3 ($^3\bar{\text{He}}$) during the following two decades^{13–16}. Recent accelerator and detector advances led to the first production of antihydrogen¹⁷ atoms in 1995 and the discovery of strange antimatter, the antihypertriton ($^3\bar{\Lambda}\text{H}$), in 2010 at RHIC at the Brookhaven National Laboratory (ref. 18 and references therein).

Collisions of relativistic heavy nuclei create suitable conditions for producing antinuclei, because large amounts of energy are deposited into a more extended volume¹⁹ than that achieved in elementary particle collisions. These nuclear interactions briefly ($\sim 10^{-23}$ s) produce hot and dense matter containing roughly equal numbers of quarks and antiquarks²⁰, often interpreted as quark gluon plasma²¹. In contrast to the Big Bang, nuclear collisions produce negligible gravitational attraction and allow the plasma to expand rapidly. The hot and dense matter cools down and undergoes a transition into a hadron gas, producing nucleons and their antiparticles. The production of light antinuclei can be modelled successfully by macroscopic thermodynamics⁷, which assumes energy equipartition, or by a microscopic coalescence process^{8,22}, which assumes uncorrelated probabilities for antinucleons close in position and momentum to become bound. The high temperature and high antibaryon density of relativistic heavy ion collisions provide a favourable environment for both production mechanisms.

The central detector used in our measurements of antimatter, the Time Projection Chamber (TPC)²³ of the STAR experiment (Solenoidal

Tracker At RHIC), is situated in a solenoidal magnetic field and is used for three-dimensional imaging of the ionization trail left along the path of charged particles (Fig. 1). In addition to the momentum provided by the track curvature in the magnetic field, the detection of $^4\bar{\text{He}}$ particles relies on two key measurements: the mean energy loss per unit track length, $\langle dE/dx \rangle$, in the TPC gas, which helps distinguish particles with different masses or charges, and the time of flight of particles arriving at the time of flight barrel (TOF)²⁴ surrounding the TPC. In general, time of flight provides particle identification in a higher momentum range than $\langle dE/dx \rangle$. The $\langle dE/dx \rangle$ resolution is 7.5% and the timing resolution for the TOF system is 95 ps within a 7–75 ns window.

The trigger system at STAR selects collisions of interest for analysis. The minimum-bias trigger selects all particle-producing collisions, regardless of the extent of overlap of the incident nuclei. A central trigger (CENT) preferentially selects head-on collisions, rejecting about 90% of the events acquired using the minimum-bias trigger. The sample of 10^9 Au+Au collisions used in this search is selected on the basis of the minimum-bias trigger, on CENT, and on various specialized triggers. Preferential selection of events containing tracks with charge $Ze = \pm 2e$ (where e is the electron charge and Z is the particle charge in units of e) was implemented using a High-Level Trigger (HLT) for data acquired in 2010. The HLT used computational resources at STAR to perform a real-time fast track reconstruction to tag events that had at least one track with a $\langle dE/dx \rangle$ value that is larger than a threshold set to three standard deviations below the theoretically expected value²⁵ for $^3\bar{\text{He}}$ at the same magnetic rigidity. The HLT successfully identified 70% of the events where a $^4\bar{\text{He}}$ track was present while selecting only 0.4% of the events for express analyses.

Figure 2 shows $\langle dE/dx \rangle$ versus the magnitude of magnetic rigidity, $p/|Z|$, where p is momentum. A distinct band of positive particles

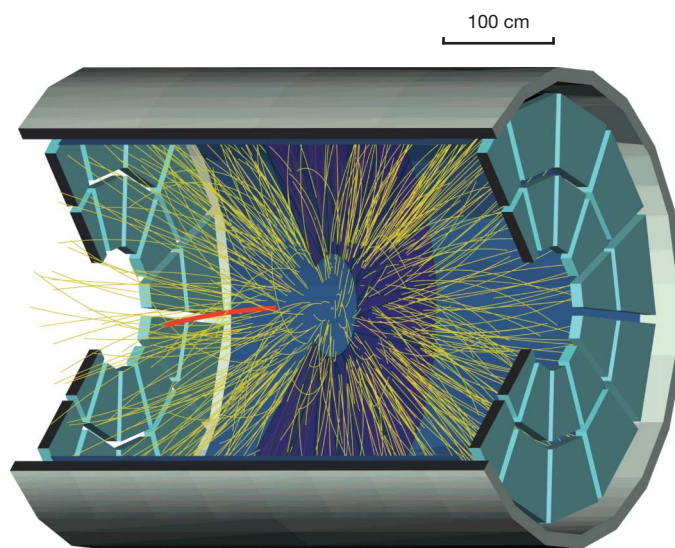


Figure 1 | A three-dimensional rendering of the STAR TPC surrounded by the TOF barrel shown as the outermost cylinder. Tracks from an event which contains a $^4\bar{\text{He}}$ are shown, with the $^4\bar{\text{He}}$ track highlighted in bold red.

*Lists of participants and their affiliations appear at the end of the paper.

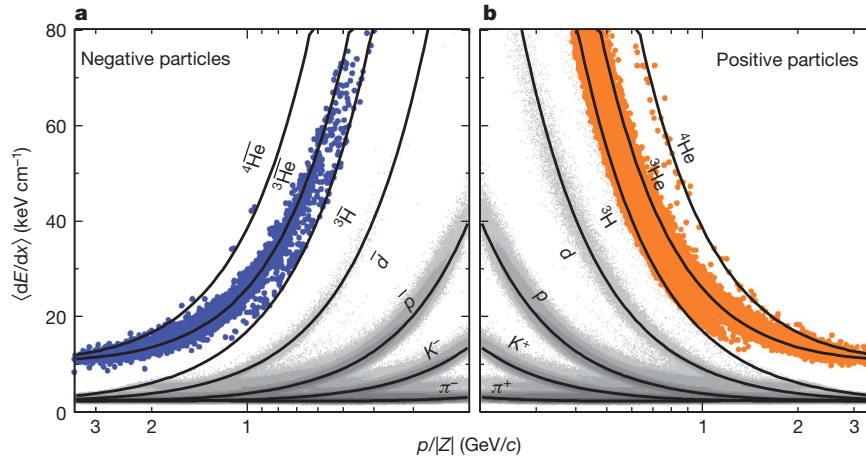


Figure 2 | $\langle dE/dx \rangle$ versus $p/|Z|$. **a**, For negatively charged particles (grey and blue dots); **b**, for positively charged particles (grey and orange dots). The black curves show the expected values for each species. The lower edges of the bands of coloured dots correspond to the online calculation by the HLT of 3σ below

centred around the expected value²⁵ for ^4He particles is shown in Fig. 2b and indicates that the detector is well-calibrated. In Fig. 2a, where $p/|Z|$ is less than 1.4 GeV/c (where c is the velocity of light), four negative particles are particularly well separated from the ^3He band and are located within the expected band for ^4He . Above 1.75 GeV/c, $\langle dE/dx \rangle$ values of ^3He and ^4He merge and the TOF system is needed to separate these two species.

Figure 3a and b shows the $\langle dE/dx \rangle$ (in units of multiples of $\sigma_{dE/dx}$, $n_{\sigma_{dE/dx}}$) versus calculated mass $m = (p/c)\sqrt{(t^2c^2/L^2 - 1)}$, where $\sigma_{dE/dx}$ is the r.m.s. width of the $\langle dE/dx \rangle$ distribution for ^4He or ^3He , and t and L are the time of flight and path length, respectively. Negatively and positively charged particles are shown in Fig. 3a and b, respectively. In both panels, majority species are ^3He and ^3He . In Fig. 3b, the ^4He particles cluster around $n_{\sigma_{dE/dx}} = 0$ and mass 3.73 GeV/ c^2 , the appropriate mass for ^4He . A similar but smaller cluster of particles can be found in Fig. 3a for ^4He . In Fig. 3c we show the projection onto the mass axis for particles in Fig. 3a and b with $n_{\sigma_{dE/dx}}$ of -2 to 3 . There is clear separation between ^3He and ^4He mass peaks. Eighteen counts for ^4He are observed. Of those, sixteen are from collisions recorded in 2010. Two counts²⁶ identified by $\langle dE/dx \rangle$ alone from data recorded in 2007 are not included in this figure, because the STAR TOF was not installed at that time.

To evaluate the background in ^4He due to ^3He contamination, we simulate the ^3He mass distribution with momenta and path lengths, as well as the expected time of flight from ^3He particles with timing resolution derived from the same data sample. The contamination from misidentifying ^3He as ^4He is estimated by integrating over the region of the ^4He selection. We estimate that the background contributes 1.4 (0.05) counts of the 15 (1) total counts from Au+Au collisions at 200 (62) GeV recorded in 2010. Therefore, the probability of misidentification is at the 10^{-11} level.

The observed counts are used to calculate the antimatter yield with appropriate normalization (the differential invariant yield) in order to compare to the theoretical expectation. Detector acceptance, efficiency, and antimatter annihilation with the detector material are taken into account when computing yields. Various uncertainties related to tracking in the TPC, matching in the TOF, and triggering in the HLT are cancelled when the yield ratios of $^4\text{He}/^3\text{He}$ and $^4\text{He}/^3\text{He}$ are calculated. The ratios are $^4\text{He}/^3\text{He} = (3.0 \pm 1.3(\text{stat})^{+0.5}_{-0.3}(\text{sys})) \times 10^{-3}$ and $^4\text{He}/^3\text{He} = (3.2 \pm 2.3(\text{stat})^{+0.7}_{-0.2}(\text{sys})) \times 10^{-3}$ for central Au+Au collisions at 200 GeV (where ‘stat’ and ‘sys’ indicate the statistical and systematic errors). The ratios were obtained in two windows. The first was $40^\circ < \theta < 140^\circ$, where the polar angle, θ , is the angle between the particle’s momentum vector and the beam axis (these θ limits correspond to limits of -1 to 1 in a related quantity, pseudorapidity). The second was a p_T per baryon window centred at $p_T/$

the $\langle dE/dx \rangle$ band centre²⁵ for ^3He . The grey bands correspond to charged particles which lie far from the region of particular focus in the present study, and which were not selected by the HLT. The bands marked p , \bar{p} , K and π correspond to protons, antiprotons, kaons and pions, respectively.

$|B| = 0.875$ GeV/c with a width of 0.25 GeV/c, where p_T is the projection of the momentum vector on the plane that is transverse to the beam axis. Ratios calculated by a Blastwave model²⁷ for the $p_T/|B|$ window mentioned above and for the whole range of $p_T/|B|$ differ by only 1%. The differential yields (see legend to Fig. 4) for ^4He (^4He) are obtained by multiplying the ratio of $^4\text{He}/^3\text{He}$ ($^4\text{He}/^3\text{He}$) with the ^3He (^3He)

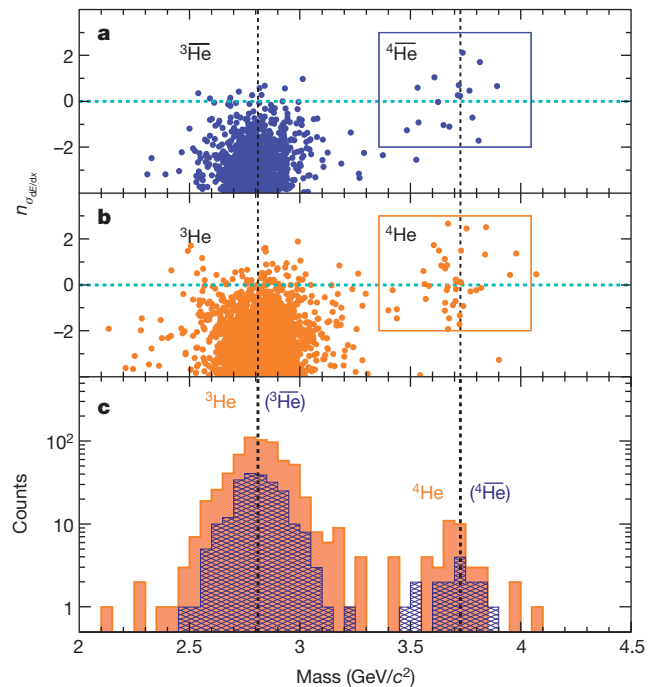


Figure 3 | Isotope identification based on energy loss and mass calculated from momentum per charge and time of flight. **a**, **b**, The $\langle dE/dx \rangle$ in units of multiples of $\sigma_{dE/dx}$, $n_{\sigma_{dE/dx}}$, of negatively charged particles (**a**) and positively charged particles (**b**) as a function of mass measured by the TOF system. The masses of ^3He (^3He) and ^4He (^4He) are indicated by the black vertical dashed lines at 2.81 GeV/ c^2 and 3.73 GeV/ c^2 , respectively. The light blue horizontal dashed line marks the position of zero deviation from the expected value of $\langle dE/dx \rangle$ ($n_{\sigma_{dE/dx}} = 0$) for ^4He (^4He). The rectangular boxes highlight areas for ^4He (^4He) selections: $-2 < n_{\sigma_{dE/dx}} < 3$ and $3.35 \text{ GeV}/c^2 < \text{mass} < 4.04 \text{ GeV}/c^2$ (corresponding to a $\pm 3\sigma$ window in mass). **c**, A projection of entries in **a** and **b** onto the mass axis for particles in the window of $-2 < n_{\sigma_{dE/dx}} < 3$. The combined measurements of energy loss and the time of flight allow a clean identification to be made in a sample of 0.5×10^{12} tracks from 10^9 Au+Au collisions.

yields²⁸. The systematic uncertainties consist of background (−6% for both ratios), feed-down from (anti-)hypertritons (18% for both ${}^3\text{He}$ and ${}^3\bar{\text{He}}$), knockouts from beam–material interactions (−5% for the ratio ${}^4\text{He}/{}^3\text{He}$ only) and absorption (4% for the ratio ${}^4\text{He}/{}^3\text{He}$ only). Figure 4 shows the exponential³ invariant yields versus baryon number in 200 GeV central Au+Au collisions. Empirically, the production rate reduces by a factor of $1.6^{+1.0}_{-0.6} \times 10^3 (1.1^{+0.3}_{-0.2} \times 10^3)$ for each additional antinucleon (nucleon) added to the antinucleus (nucleus). This general trend is expected from coalescent nucleosynthesis models⁸, originally developed to describe production of antideuterons²², as well as from thermodynamic models⁷.

In a microscopic picture, a light nucleus emerging from a relativistic heavy-ion collision is produced during the last stage of the collision process. The quantum wavefunctions of the constituent nucleons, if close enough in momentum and coordinate space, will overlap to produce the nucleus. The production rate for a nucleus with baryon number B is proportional to the nucleon density in momentum and coordinate space, raised to the power of $|B|$, and therefore exhibits exponential behaviour as a function of B . Alternatively, in a thermodynamic model, a nucleus is regarded as an object with energy $E \approx |B|m_N$, where m_N is the nucleon mass, and the production rate is determined by the Boltzmann factor $\exp(-E/T)$, where T is the temperature^{3,7}. This model also produces an exponential yield. A more rigorous calculation⁵ can provide a good fit to the available particle yields, and predicts the ratios integrated over p_T to be ${}^4\text{He}/{}^3\text{He} = 3.1 \times 10^{-3}$ and ${}^4\bar{\text{He}}/{}^3\bar{\text{He}} = 2.4 \times 10^{-3}$, consistent with our measurements. The considerations outlined above offer a good estimate for the production rate of even heavier antinuclei. For example, the yield of the stable antimatter nucleus next in line ($B = -6$) is predicted to be down by a factor of 2.6×10^6 compared to ${}^4\bar{\text{He}}$ and is beyond the reach of current accelerator technology.

A potentially more copious production mechanism for heavier antimatter is by the direct excitation of complex nuclear structures from the vacuum²⁹. A deviation from the usual rate reduction with increasing mass would be an indication of a radically new production mechanism⁷. On the other hand, going beyond nuclear physics, the sensitivity of current and planned space-based charged particle detectors is below what would be needed to observe antihelium produced by nuclear interactions in the cosmos, and consequently, any observation of antihelium or even heavier antinuclei in space would indicate the

existence of a large amount of antimatter elsewhere in the Universe. In particular, finding ${}^4\bar{\text{He}}$ in the cosmos is one of the major motivations for space detectors such as the Alpha Magnetic Spectrometer³⁰. We have shown that ${}^4\bar{\text{He}}$ exists, and have measured its rate of production in nuclear interactions, providing a point of reference for possible future observations in cosmic radiation. Barring one of those dramatic discoveries mentioned above or a new breakthrough in accelerator technology, it is likely that ${}^4\bar{\text{He}}$ will remain the heaviest stable antimatter nucleus observed for the foreseeable future.

Received 14 March; accepted 4 April 2011.

Published online 24 April 2011.

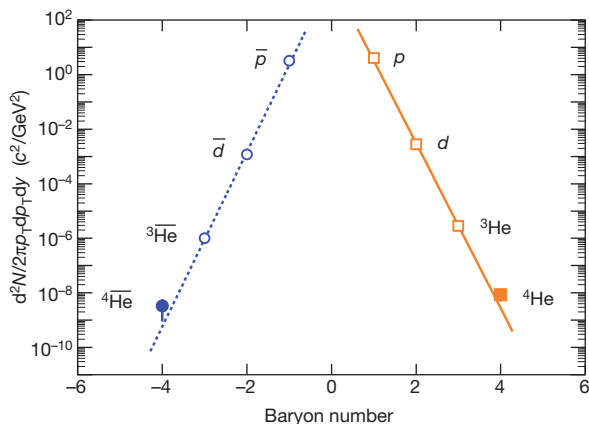


Figure 4 | Differential invariant yields as a function of baryon number, B . The differential invariant yields $d^2N/(2\pi p_T dp_T dy)$ were evaluated at $p_T/|B| = 0.875$ GeV/c, in central 200 GeV Au+Au collisions, where N is counts per event and y is rapidity. Yields for (anti)tritons (${}^3\text{H}$ and ${}^3\bar{\text{H}}$) lie close to the positions for ${}^3\text{He}$ and ${}^3\bar{\text{He}}$, but are not included here because of poorer identification of (anti)tritons. The lines represent fits with the exponential formula $\propto e^{-r|B|}$ for positive (solid orange line) and negative (dashed blue line) particles separately, where r is the production reduction factor. Analysis details of yields other than ${}^4\text{He}$ (${}^4\bar{\text{He}}$) have been presented elsewhere^{4,28} and are plotted here as open symbols. The plotted error bars show standard statistical errors only. Systematic errors are smaller than the symbol size, and are not plotted.

- Lemaître, G. A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extragalactic nebulae [in French]. *Ann. Soc. Sci. Brux.* **47**, 49–55 (1927).
- Wiebel-Sooth, B., Biermann, P. & Meyer, H. Cosmic rays, VII. Individual element spectra: prediction and data. *Astron. Astrophys.* **330**, 389–398 (1998).
- Armstrong, T. A. *et al.* Mass dependence of light-nucleus production in ultrarelativistic heavy-ion collisions. *Phys. Rev. Lett.* **83**, 5431–5434 (1999).
- Liu, H. *Production of Meson, Baryon and Light Nuclei ($A=2,3$): Investigating Freeze-Out Dynamics and Roles of Energetic Quarks and Gluons in Au+Au Collisions at RHIC*. PhD thesis, Univ. Sci. Technol. China (2007).
- Andronic, A., Braun-Munzinger, P., Stachel, J. & Stöcker, H. Production of light nuclei, hypernuclei and their antiparticles in relativistic nuclear collisions. *Phys. Lett. B* **697**, 203–207 (2011).
- Harrison, M., Ludlam, T. & Ozaki, S. The relativistic heavy ion collider project: RHIC and its detectors. *Nucl. Instrum. Methods Phys. Res. A* **499**, 235–244 (2003).
- Braun-Munzinger, P. & Stachel, J. The quest for the quark-gluon plasma. *Nature* **448**, 302–309 (2007).
- Sato, H. & Yazaki, K. On the coalescence model for high energy nuclear reactions. *Phys. Lett. B* **98**, 153–157 (1981).
- Dirac, P. A. M. The quantum theory of the electron. *Proc. R. Soc. Lond. A* **117**, 610–624 (1928).
- Anderson, C. D. The positive electron. *Phys. Rev.* **43**, 491–494 (1933).
- Chamberlain, O., Segrè, E., Wiegand, C. & Ypsilantis, T. Observation of antiprotons. *Phys. Rev.* **100**, 947–950 (1955).
- Cork, B., Lambertson, G. R., Piccioni, O. & Wenzel, W. A. Antineutrons produced from antiprotons in charge-exchange collisions. *Phys. Rev.* **104**, 1193–1197 (1956).
- Massam, T., Müller, T., Righini, B., Schneegans, M. & Zichichi, A. Experimental observation of antideuteron production. *Nuovo Cim.* **39**, 10–14 (1965).
- Dorfan, D. E., Eades, J., Lederman, L. M., Lee, W. & Ting, C. C. Observation of antideuterons. *Phys. Rev. Lett.* **14**, 1003–1006 (1965).
- Vishnevsky, N. K. *et al.* Observation of antitritium [in Russian]. *Yad. Fiz.* **20**, 694–708 (1974).
- Antipov, Y. M. *et al.* Observation of antihelium3 [in Russian]. *Yad. Fiz.* **12**, 311–322 (1970).
- Baur, G. *et al.* Production of antihydrogen. *Phys. Lett. B* **368**, 251–258 (1996).
- Abelev, B. I. *et al.* Observation of an antimatter hypernucleus. *Science* **328**, 58–62 (2010).
- Lee, T. D. Abnormal nuclear states and vacuum excitation. *Rev. Mod. Phys.* **47**, 267–275 (1975).
- BRAHMS, PHENIX, PHOBOS, and STAR Collaboration. RHIC white papers. *Nucl. Phys. A* **757**, Issues 1–2, 1–283 (2005).
- Heinz, U., Subramanian, P. R., Stöcker, H. & Greiner, W. Formation of antimatter clusters in the hadronization phase transition. *J. Phys. G* **12**, 1237–1263 (1986).
- Butler, S. T. & Pearson, C. A. Deuterons from high-energy proton bombardment of matter. *Phys. Rev. Lett.* **7**, 69–71 (1961).
- Anderson, M. *et al.* The STAR time projection chamber: a unique tool for studying high multiplicity events at RHIC. *Nucl. Instrum. Methods Phys. Res. A* **499**, 659–678 (2003).
- Bonner, B. *et al.* A single time-of-flight tray based on multigap resistive plate chambers for the STAR experiment at RHIC. *Nucl. Instrum. Methods Phys. Res. A* **508**, 181–184 (2003).
- Bichsel, H. A method to improve tracking and particle identification in TPCs and silicon detectors. *Nucl. Instrum. Methods Phys. Res. A* **562**, 154–197 (2006).
- Zhou, J. *Light (Anti)Nuclei Production in the STAR Experiment at RHIC*. Ph.D. thesis, Rice Univ. (2009).
- Tang, Z. *et al.* Spectra and radial flow in relativistic heavy ion collisions with Tsallis statistics in a blast-wave description. *Phys. Rev. C* **79**, 051901(R) (2009).
- Abelev, B. I. *et al.*, STAR Collaboration. Yields and elliptic flow of d (\bar{d}) and ${}^3\text{He}$ (${}^3\bar{\text{He}}$) in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV. Preprint at (<http://arxiv.org/abs/0909.0566>) (2009).
- Greiner, W. On the extension of the periodic system into the sectors of strangeness and antimatter. *Int. J. Mod. Phys. E* **5**, 1–90 (1996).
- Ahlen, S. *et al.* An antimatter spectrometer in space. *Nucl. Instrum. Methods Phys. Res. A* **350**, 351–367 (1994).

Acknowledgements We thank the RHIC Operations Group and RACF at BNL, the NERSC Center at LBNL and the Open Science Grid consortium for providing resources and support. This work was supported in part by the Offices of NP and HEP within the US DOE Office of Science, the US NSF, the Sloan Foundation, the DFG cluster of excellence ‘Origin and Structure of the Universe’ of Germany, CNRS/IN2P3, FAPESP

CNPq of Brazil, the Ministry of Education and Science of the Russian Federation, NNSFC, CAS, MoST and MoE of China, GA and MSMT of the Czech Republic, FOM and NWO of the Netherlands, DAE, DST and CSIR of India, the Polish Ministry of Science and Higher Education, the Korea Research Foundation, the Ministry of Science, Education and Sports of Croatia, and RosAtom of Russia.

Author Contributions All authors contributed equally.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to The Star Collaboration (star-antihe4-l@lists.bnl.gov).

The STAR Collaboration

H. Agakishiev¹, M. M. Aggarwal², Z. Ahammed³, A. V. Alakhverdyants¹, I. Alekseev⁴, J. Alford⁵, B. D. Anderson⁵, C. D. Anson⁶, D. Arkhipkin⁷, G. S. Averichev¹, J. Balewski⁸, D. R. Beavis⁹, N. K. Behera³, R. Bellwied¹⁰, M. J. Betancourt⁵, R. R. Betts¹¹, A. Bhasin¹², A. K. Bhati², H. Bichsel¹³, J. Bielcik¹⁴, J. Bielcikova¹⁵, B. Biritz¹⁶, L. C. Bland⁷, I. G. Bordyuzhin⁴, W. Borowski¹⁷, J. Bouchet¹⁸, E. Braidot¹⁸, A. V. Brandin¹⁹, A. Bridgeman²⁰, S. G. Brovko²¹, E. Bruna²², S. Bueltmann²³, I. Bunzarov¹, T. P. Burton⁷, X. Z. Cai²⁴, H. Caines²², M. Calderon²¹, D. Cebra²¹, R. Cendejas¹⁶, M. C. Cervantes²⁵, Z. Chajec²⁶, P. Chaloupka¹⁵, S. Chattopadhyay²⁶, H. F. Chen²⁷, J. H. Chen²⁴, J. Y. Chen²⁸, L. Chen²⁸, J. Cheng²⁹, M. Cherney³⁰, A. Chikanian²², K. E. Choi³¹, W. Christie⁷, P. Chung¹⁵, M. J. M. Codrington²⁵, R. Corliss⁸, J. G. Cramer¹³, H. J. Crawford³², A. Davila Leyva³³, L. C. De Silva¹⁰, R. R. Debebe⁷, T. G. Dedovich¹, A. A. Derevschikov³⁴, R. Derradi de Souza³⁵, L. Didenko⁷, P. Djawotho²⁵, S. M. Dogra¹², X. Dong³, J. L. Drachenberg²⁵, J. E. Draper²¹, J. C. Dunlop⁷, L. G. Efimov¹, M. Elinim³⁶, J. Engelage³², G. Eppley³⁷, M. Estienne¹⁷, L. Eun³⁸, O. Evdokimov¹¹, R. Fatemi³⁹, J. Fedorisin¹, R. G. Fersch³⁹, P. Filip¹, E. Finch²⁴, V. Fine⁷, Y. Fisyak⁷, C. A. Gagliardi²⁵, D. R. Gangadharan¹⁶, F. Geurts³⁷, P. Ghosh²⁶, Y. N. Gorbunov³⁰, A. Gordon⁷, O. G. Grebenyuk³, D. Grosnick⁴⁰, S. M. Guertin¹⁶, A. Gupta¹², S. Gupta¹², W. Gurny⁷, B. Haag⁷, O. Hajkova¹⁴, A. Hamed²⁵, L.-X. Han²⁴, J. W. Harris²², J. P. Hays-Wehle⁸, M. Heinz²², S. Heppelmann³⁸, A. Hirsch⁴¹, E. Hjort³, G. W. Hoffmann³³, D. J. Hofman¹¹, B. Huang²⁷, H. Z. Huang¹⁶, T. J. Humanic⁶, L. Huo²⁵, G. Igo¹⁶, P. Jacobs³, W. W. Jacobs⁴², C. Jena⁴³, F. Jin²⁴, J. Joseph⁵, E. G. Judd³², S. Kabana¹⁷, K. Kang²⁹, J. Kapitan¹⁵, K. Kauder¹¹, H. W. Ke²⁸, D. Keane⁵, A. Kechechyan¹, D. Kettler¹³, D. P. Kikola⁴¹, J. Kiryluk³, A. Kisiel⁴⁴, V. Kizka¹, S. R. Klein³, A. G. Knospe²², D. D. Koetke⁴⁰, T. Kollegger⁴⁵, J. Konzer⁴¹, I. Koralt²³, L. Koroleva⁴, W. Korsch³⁹, L. Kotchenda¹⁹, V. Kouchpil¹⁵, P. Kravtsov¹⁹, K. Krueger²⁰, M. Krus¹⁴, L. Kumar⁵, P. Kurnadi¹⁶, M. A. C. Lamont⁷, J. M. Landgraf⁷, S. LaPointe³⁶, J. Lauret⁷, A. Lebedev⁷, R. Lednicky¹, J. H. Lee⁷, W. Light⁹, M. J. LeVine³, C. Li²⁷, L. Li³³, N. Li²⁸, W. Li²⁴, X. Li⁴¹, X. Li⁴⁶, Y. Li²⁹, Z. M. Li²⁸, M. A. Lisa⁶, F. Liu²⁸, H. Liu²¹, J. Liu³⁷, T. Ljubicic⁷, W. J. Llope³⁷, R. S. Longacre⁷, W. A. Love⁷, Y. Lu²⁷, E. V. Lukashov¹⁹, X. Luo²⁷, G. L. Ma²⁴, Y. G. Ma²⁴, D. P. Mahapatra⁴³, R. Majka²², O. I. Mall²¹, R. Manweiler⁴⁰, S. Margetis⁵, C. Markert³, H. Masui³, H. S. Matis⁵, Yu. A. Matulenko³⁴, D. McDonald³⁷, T. S. McShane³⁰, A. Meschanin³⁴, R. Milner⁸, N. G. Minaev³⁴, S. Mioduszewski²⁵, M. K. Mitrović⁷, Y. Mohammed²⁵, B. Mohanty²⁶, M. M. Mondal²⁶, B. Morozov⁴, D. A. Morozov³⁴, M. G. Munhoz⁴⁷, M. K. Mustafa⁴¹, M. Naglis³, B. K. Nandi⁹, T. K. Nayak²⁶, P. K. Netrakanti⁴¹, L. V. Nogach³⁴, S. B. Nurushev³⁴, G. Odyniec³, A. Ogawa⁷, K. Oh³¹, A. Ohlson²², V. Okorokov¹⁹, E. W. Oldag³³, D. Olson³, M. Pachr¹⁴, B. S. Page⁴², S. K. Pal²⁶, Y. Pandit⁵, Y. Panebratsev¹, T. Pawlak⁴⁴, H. Pei¹¹, T. Peitzmann¹⁸, C. Perkins³², W. Periyat⁴⁴, P. Pile⁷, M. Planinic⁴⁸, M. A. Ploskon³, J. Pluta⁴⁴, D. Plyku²³, N. Poljak⁴⁸, J. Porter³, A. M. Poskanzer³, B. V. K. S. Potukuchi¹², C. B. Powell³, D. Prindle¹³, C. Pruneau³⁶, N. K. Pruthi², P. R. Pujahari⁹, J. Putschke²², H. Qiu⁴⁹, R. Raniwala⁵⁰, S. Raniwala⁵⁰, R. L. Ray³³, R. Redwine³, R. Reed²¹, H. G. Ritter³, J. B. Roberts³⁷, O. V. Rogachevskiy¹, J. L. Romero²¹, L. Ruan⁷, J. Rusnak¹⁵, N. R. Sahoo²⁶, I. Sakrejda³, S. Salur²¹, J. Sandweiss²², E. Sangaline²¹, A. Sarkar³, J. Schambach³³, R. P.

Scharenberg⁴¹, A. M. Schmah³, N. Schmitz⁵¹, T. R. Schuster⁴⁵, J. Seele⁸, J. Seger³⁰, I. Selyuzhenkov⁴², P. Seyboth⁵¹, N. Shah¹⁶, E. Shahaliev¹, M. Shao²⁷, M. Sharma³⁶, S. S. Shi²⁸, Q. Y. Shou²⁴, E. P. Sichtermann³, F. Simon⁵¹, R. N. Singaraju²⁶, M. J. Skoby⁴¹, N. Smirnov²², D. Solanki⁵⁰, P. Sorensen⁷, H. M. Spinka²⁰, B. Srivastava⁴¹, T. D. S. Stanislaus⁴⁰, D. Staszak¹⁶, S. G. Steadman⁸, J. R. Stevens⁴², R. Stock⁴⁵, M. Strikhanov¹⁹, B. Stringfellow⁴¹, A. A. P. Suaide⁴⁷, M. C. Suarez¹¹, N. L. Subba⁵, M. Sumera¹⁵, X. M. Sun³, Y. Sun²⁷, Z. Sun⁴⁹, B. Surrow⁸, D. N. Svirida⁴, T. J. M. Symons³, A. Szanto de Toledo⁴⁷, J. Takahashi³⁵, A. H. Tang⁷, Z. Tang²⁷, L. H. Tarini³⁶, T. Tarnowsky⁵², D. Thein³³, J. H. Thomas³, J. Tian²⁴, A. R. Timmins¹⁰, D. Tlustý¹⁵, M. Tokarev¹, T. A. Trainor¹³, S. Trentalange¹⁶, R. E. Tribble²⁵, P. Tribedy²⁶, O. D. Tsai¹⁶, T. Ullrich⁷, D. G. Underwood²⁰, G. Van Buren⁷, G. van Nieuwenhuizen⁸, J. A. Vanfossen Jr⁵, R. Varma⁹, G. M. S. Vasconcelos³⁵, A. N. Vasiliev³⁴, F. Videbaek⁷, Y. P. Viyogi²⁶, S. Vokal¹, S. A. Voloshin³⁶, M. Wada³³, M. Walker⁸, F. Wang⁴¹, G. Wang¹⁶, H. Wang⁵², J. S. Wang⁴⁹, Q. Wang⁴¹, X. L. Wang²⁹, Y. Wang²⁹, G. Webb³⁹, J. C. Webb⁷, G. D. Westfall⁵², C. Whitten Jr¹⁶, H. Wieman³, S. W. Wissink⁴², R. Witt⁵³, W. Witzke³⁹, Y. F. Wu²⁸, Z. Xiao²⁹, W. Xie⁴¹, H. Xu⁴⁹, N. Xu³²⁸, Q. H. Xu⁴⁶, W. Xu¹⁶, Y. Xu²⁷, Z. Xu⁷, L. Xue²⁴, Y. Yang⁴⁹, Y. Yang²⁸, P. Yepes³⁷, K. Yip⁷, I.-K. Yoo³¹, M. Zawisza⁴⁴, H. Zbroszczyk⁴⁴, W. Zhan⁴⁹, J. B. Zhang²⁸, S. Zhang²⁴, W. M. Zhang⁵, X. P. Zhang²⁹, Y. Zhang³, Z. P. Zhang²⁷, F. Zhao¹⁶, J. Zhao²⁴, C. Zhong²⁴, W. Zhou⁴⁶, X. Zhu²⁹, Y. H. Zhu²⁴, R. Zoulkarneev¹ & Y. Zoulkarneeva¹.

¹Joint Institute for Nuclear Research, Dubna, 141 980, Russia. ²Panjab University, Chandigarh 160014, India. ³Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁴Alkhanov Institute for Theoretical and Experimental Physics, Moscow 117218, Russia. ⁵Kent State University, Kent, Ohio 44242, USA. ⁶Ohio State University, Columbus, Ohio 43210, USA. ⁷Brookhaven National Laboratory, Upton, New York 11973, USA. ⁸Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA. ⁹Indian Institute of Technology, Mumbai 400076, India. ¹⁰University of Houston, Houston, Texas 77204, USA. ¹¹University of Illinois at Chicago, Chicago, Illinois 60607, USA. ¹²University of Jammu, Jammu 180001, India. ¹³University of Washington, Seattle, Washington 98195, USA. ¹⁴Czech Technical University in Prague, FNSPE, Prague, 115 19, Czech Republic. ¹⁵Nuclear Physics Institute AS CR, 250 68 Rež/Prague, Czech Republic. ¹⁶University of California, Los Angeles, California 90095, USA. ¹⁷SUBATECH, Nantes 44307, France. ¹⁸NIKHEF and Utrecht University, Amsterdam 1098 XG, The Netherlands. ¹⁹Moscow Engineering Physics Institute, Moscow 115409, Russia. ²⁰Argonne National Laboratory, Argonne, Illinois 60439, USA. ²¹University of California, Davis, California 95616, USA. ²²Yale University, New Haven, Connecticut 06520, USA. ²³Old Dominion University, Norfolk, Virginia 23529, USA. ²⁴Shanghai Institute of Applied Physics, Shanghai 201800, China. ²⁵Texas A&M University, College Station, Texas 77843, USA. ²⁶Variable Energy Cyclotron Centre, Kolkata 700064, India. ²⁷University of Science & Technology of China, Hefei 230026, China. ²⁸Institute of Particle Physics, CCNU (HZNU), Wuhan 430079, China. ²⁹Tsinghua University, Beijing 100084, China. ³⁰Creighton University, Omaha, Nebraska 68178, USA. ³¹Pusan National University, Pusan 609-735, South Korea. ³²University of California, Berkeley, California 94720, USA. ³³University of Texas, Austin, Texas 78712, USA. ³⁴Institute of High Energy Physics, Protvino 142281, Russia. ³⁵Universidade Estadual de Campinas, Sao Paulo 13083-859, Brazil. ³⁶Wayne State University, Detroit, Michigan 48201, USA. ³⁷Rice University, Houston, Texas 77251, USA. ³⁸Pennsylvania State University, University Park, Pennsylvania 16802, USA. ³⁹University of Kentucky, Lexington, Kentucky, 40506-0055, USA. ⁴⁰Valparaíso University, Valparaíso, Indiana 46383, USA. ⁴¹Purdue University, West Lafayette, Indiana 47907, USA. ⁴²Indiana University, Bloomington, Indiana 47408, USA. ⁴³Institute of Physics, Bhubaneswar 751005, India. ⁴⁴Warsaw University of Technology, Warsaw 00-661, Poland. ⁴⁵University of Frankfurt, Frankfurt 60325, Germany. ⁴⁶Shandong University, Jinan, Shandong 250100, China. ⁴⁷Universidade de Sao Paulo, Sao Paulo 05508-090, Brazil. ⁴⁸University of Zagreb, Zagreb, HR-10002, Croatia. ⁴⁹Institute of Modern Physics, Lanzhou 730000, China. ⁵⁰University of Rajasthan, Jaipur 302004, India. ⁵¹Max-Planck-Institut für Physik, Munich 80805, Germany. ⁵²Michigan State University, East Lansing, Michigan 48824, USA. ⁵³United States Naval Academy, Annapolis, Maryland 21402, USA.

‡Deceased

Sharply increased mass loss from glaciers and ice caps in the Canadian Arctic Archipelago

Alex S. Gardner^{1,2}, Geir Moholdt^{3,4}, Bert Wouters⁵, Gabriel J. Wolken⁶, David O. Burgess⁷, Martin J. Sharp¹, J. Graham Cogley⁸, Carsten Braun⁹ & Claude Labine¹⁰

Mountain glaciers and ice caps are contributing significantly to present rates of sea level rise and will continue to do so over the next century and beyond^{1–5}. The Canadian Arctic Archipelago, located off the northwestern shore of Greenland, contains one-third of the global volume of land ice outside the ice sheets⁶, but its contribution to sea-level change remains largely unknown. Here we show that the Canadian Arctic Archipelago has recently lost 61 ± 7 gigatonnes per year (Gt yr^{-1}) of ice, contributing $0.17 \pm 0.02 \text{ mm yr}^{-1}$ to sea-level rise. Our estimates are of regional mass changes for the ice caps and glaciers of the Canadian Arctic Archipelago referring to the years 2004 to 2009 and are based on three independent approaches: surface mass-budget modelling plus an estimate of ice discharge ($\text{SMB} + D$), repeat satellite laser altimetry (ICESat) and repeat satellite gravimetry (GRACE). All three approaches show consistent and large mass-loss estimates. Between the periods 2004–2006 and 2007–2009, the rate of mass loss sharply increased from $31 \pm 8 \text{ Gt yr}^{-1}$ to $92 \pm 12 \text{ Gt yr}^{-1}$ in direct response to warmer summer temperatures, to which rates of ice loss are highly sensitive ($64 \pm 14 \text{ Gt yr}^{-1}$ per 1 K increase). The duration of the study is too short to establish a long-term trend, but for 2007–2009, the increase in the rate of mass loss makes the Canadian Arctic Archipelago the single largest contributor to eustatic sea-level rise outside Greenland and Antarctica.

Several long-term records (about 50 years) of the surface mass budget (surface accumulation minus surface ablation) of individual glaciers and ice caps exist for the Canadian Arctic Archipelago (CAA, see Fig. 1)^{7,8}, but extrapolation of these records to estimate the mass budget of the entire region introduces a large uncertainty. Repeat airborne laser altimetry surveys have been used to estimate that the glaciers of the CAA lost 23 Gt yr^{-1} of ice between spring 1995 and spring 2000 (ref. 9). This represents 0.063 mm yr^{-1} of sea-level rise if we take the global area of the ocean to be $362.5 \times 10^6 \text{ km}^2$ (ref. 10). Since 2000 the CAA has experienced some of the warmest summer temperatures on record, with four of the five warmest years since 1960 occurring after 2004 (Supplementary Information). Between 2005 and 2009 all CAA glaciers with long-term monitoring programmes^{7,8} experienced their most negative five-year period of surface mass budget since measurements began in the early 1960s. Here we present three independent estimates of change in total glacier mass between autumn 2003 and autumn 2009 for the northern CAA (Fig. 1; area $106,400 \text{ km}^2$) and two independent estimates for the southern CAA (Fig. 1; area $42,000 \text{ km}^2$).

The first estimate is derived using a numerical model that simulates the regional mass change resulting from the surface mass budget. Ice discharge due to the calving of icebergs from glaciers that terminate in the sea, denoted D , is added to the surface mass-budget model results to account for the total regional ice loss (model $\text{SMB} + D$) (Supplementary Information). The model is not applied to the southern CAA because there are too few records of glacier mass budget and near-surface temperature with which to calibrate the model. The second

estimate derives mass change from the change in land-ice volume measured using repeat laser altimetry from the Ice, Cloud and Land Elevation Satellite (ICESat)¹¹. The third estimate is derived using repeat gravity observations collected by the Gravity Recovery and Climate Experiment (GRACE) satellites. The three methods are independent and produce consistent estimates of changes in glacier mass for the years 2004 to 2009 (Fig. 2), where each year refers to the mass-budget year starting in the autumn of the previous calendar year. All estimates are given as the mean $\pm 2\sigma$ (95% confidence interval).

In general, the CAA receives low amounts of precipitation ($100\text{--}300 \text{ kg m}^{-2} \text{ yr}^{-1}$) with locally higher rates ($300\text{--}1,000 \text{ kg m}^{-2} \text{ yr}^{-1}$)

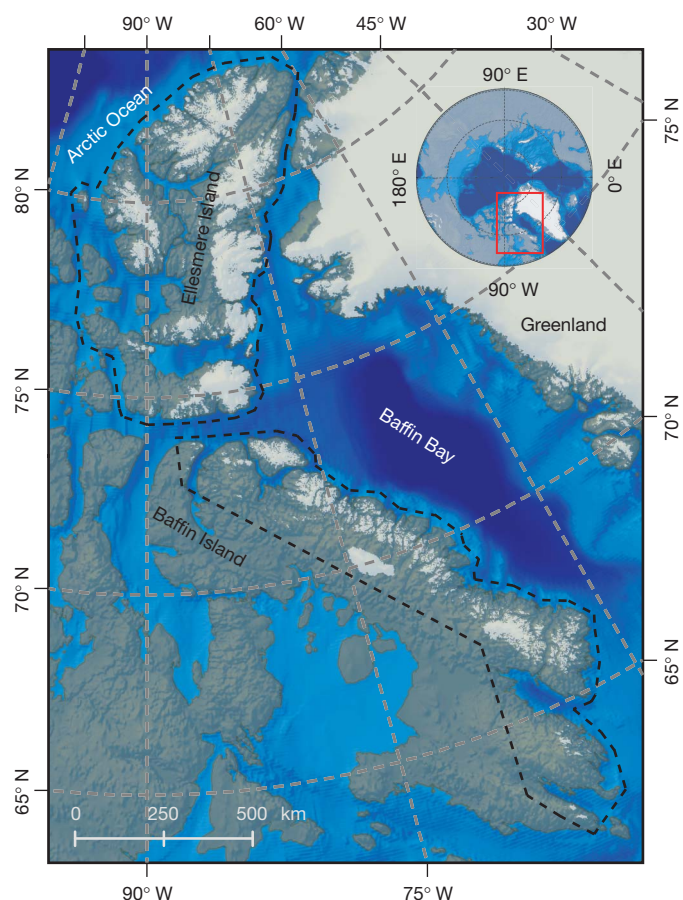


Figure 1 | Glaciers and ice caps of the Canadian Arctic Archipelago. Black dashed lines delineate the northern and southern study regions. The main panel is an enlargement of the red rectangle superimposed on the map of the Arctic (inset).

¹Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, T6G 2E3, Canada. ²Department of Atmospheric, Oceanic and Space Science, University of Michigan, Ann Arbor, Michigan 48109, USA. ³Department of Geosciences, University of Oslo, N-0316 Oslo, Norway. ⁴Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, La Jolla, California 92093, USA. ⁵The Royal Netherlands Meteorological Institute, NL-3730 AE De Bilt, Netherlands. ⁶Division of Geological and Geophysical Surveys, Alaska Department of Natural Resources, Fairbanks, Alaska 99709, USA. ⁷Geological Survey of Canada, Ottawa, Ontario, K1A 0E8, Canada. ⁸Department of Geography, Trent University, Peterborough, Ontario, K9J 7B8, Canada. ⁹Department of Geography and Regional Planning, Westfield State University, Westfield, Massachusetts 01086, USA. ¹⁰Campbell Scientific Canada Corp., Edmonton, Alberta, T5M 1W7, Canada.

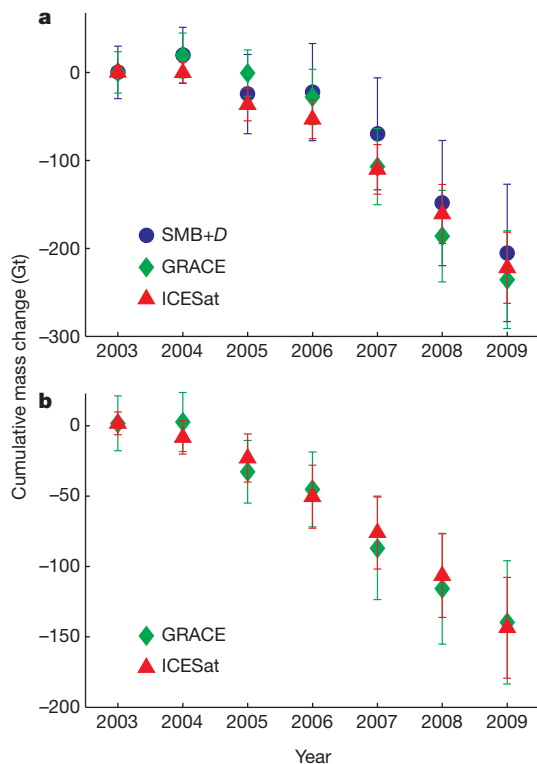


Figure 2 | Cumulative change in glacier mass between autumn 2003 and autumn 2009. Separate estimates are provided for the northern (a) and southern (b) CAA. Error bars represent the 95% confidence interval.

concentrated on the east-facing slopes flanking Baffin Bay (Fig. 1). Surface air temperatures over ice masses in the region exceeded the freezing point during only two to three months of the year. Because there is generally low interannual variability in precipitation and high variability in melt production, interannual variability in the regional surface mass budget is largely governed by changes in the summer surface energy budget⁷. These are strongly correlated with summer surface air temperatures^{12–14}, which are, in turn, highly dependent on local synoptic conditions^{15,16}. In this study we apply a surface mass-budget model that determines surface melt using the temperature-index method^{17,18}. The model is forced with downscaled¹⁹ and bias-corrected temperature and precipitation fields from the National Centers for Environmental Prediction/National Center for Atmospheric Research reanalysis (Supplementary Information). For the years 2004 to 2009 the modelled mass loss from the surface mass budget (SMB) plus ice discharge (D), where $D = 4.6 \pm 1.9 \text{ Gt yr}^{-1}$ (Supplementary Information), of the northern CAA was $34 \pm 13 \text{ Gt yr}^{-1}$ (Fig. 3). The average mass loss from the northern CAA was $7 \pm 18 \text{ Gt yr}^{-1}$ for the years 2004 to 2006, increasing to $61 \pm 18 \text{ Gt yr}^{-1}$ for the years 2007 to 2009 with a peak loss of $79 \pm 30 \text{ Gt yr}^{-1}$ in 2008. The difference between the two periods is primarily due to a 42 Gt yr^{-1} increase in melt production, which resulted from regionally warmer summer air temperatures in the lower troposphere. Warmer temperatures also contributed to a 7% decrease in snow fraction. A slight decrease in annual precipitation amount, and changes in the amount of meltwater retained by the annual snowpack, contributed another 12 Gt yr^{-1} to the increased mass loss.

For both the northern and southern CAA, we derived elevation changes from ICESat's Geoscience Laser Altimeter System (GLAS) for the period 2003–2009 (ref. 20). Elevation changes are estimated relative to rectangular planes that are fitted to 700-m-long segments of near-repeat-track data²¹. The planes represent a simplified surface topography such that multi-temporal elevation measurements that are slightly offset in location can be compared. We then extrapolate elevation changes to volume changes and convert them to mass

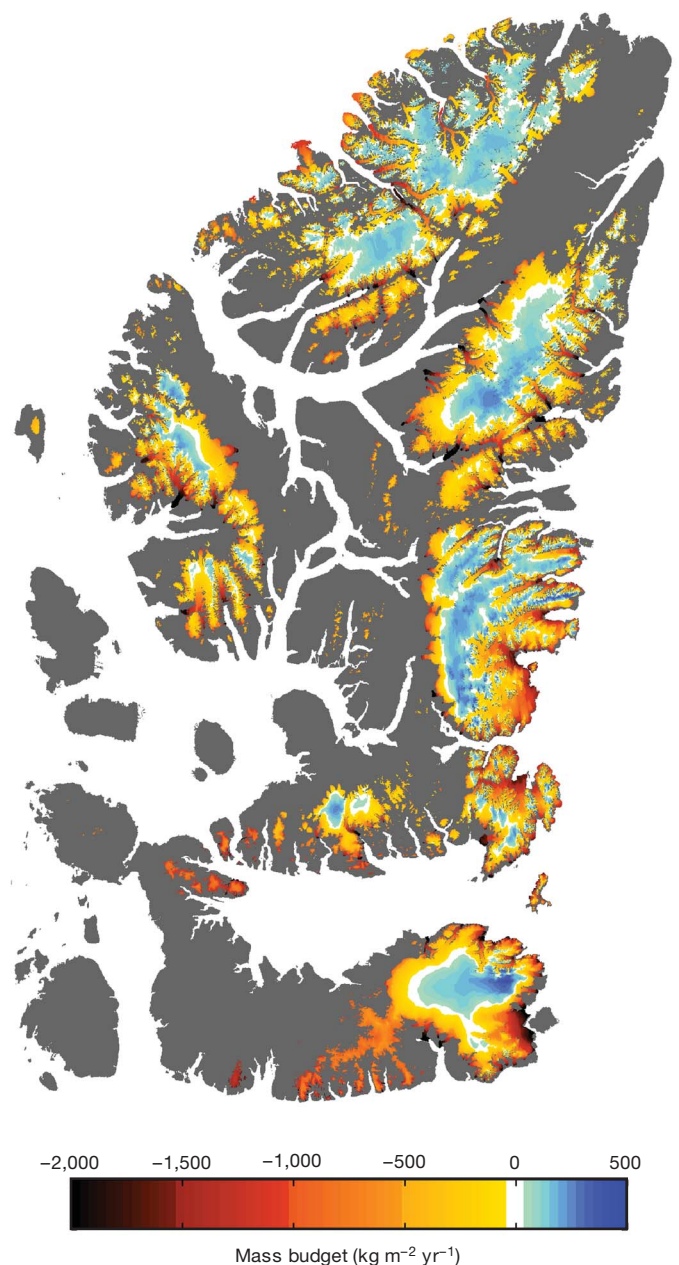


Figure 3 | Modelled surface mass budget of the northern CAA between autumn 2003 and autumn 2009. The model resolution of 0.5 km allows us to resolve the highly negative surface mass budgets of the outlet-glacier tongues.

changes using a plausible range of firn and ice densities (Supplementary Information). For the years 2004 to 2009, ICESat results show that the northern CAA lost $37 \pm 7 \text{ Gt yr}^{-1}$ and that the southern CAA lost $24 \pm 6 \text{ Gt yr}^{-1}$. ICESat results show increases in mass loss between 2004–2006 and 2007–2009 of 39 Gt yr^{-1} and 14 Gt yr^{-1} for the northern and southern CAA, respectively. Recent observations in both Alaska²² and Greenland²³ have found that marine-terminating glaciers are thinning more rapidly than land-terminating glaciers. To assess whether the same phenomenon is occurring in the CAA, we separately determined elevation changes for marine- and land-terminating glacier basins (Supplementary Information). Our results show no difference in the area-averaged rate of elevation change between the two basin types, suggesting that total ice discharge from marine-terminating glaciers has not accelerated in recent years. This gives increased confidence in both the extrapolation of ICESat elevation changes and our estimate of ice discharge.

Lastly, we derived mass changes for both the northern and southern CAA from GRACE gravity measurements. Mass-change estimates from GRACE agree very well with the other two data sets for the northern CAA, with an average mass loss between 2004 and 2009 of $39 \pm 9 \text{ Gt yr}^{-1}$. The observations confirm the sharp increase in northern CAA mass loss between 2004–2006 and 2007–2009, with an increase in the average mass loss of 60 Gt yr^{-1} . The southern CAA is estimated to have lost ice at an average rate of $24 \pm 7 \text{ Gt yr}^{-1}$ over the six-year study period, with a 16 Gt yr^{-1} increase in the rate of loss between the first three and last three years, and is in very good agreement with ICESat. The most likely sources of the disagreement between the three methods are: uncertainties in constraining the terrestrial water storage in the GRACE estimates, the identification of the appropriate end-of-season mass change in the GRACE signal, and fewer ICESat elevation retrievals in 2009 (Supplementary Information).

The error-weighted mean of all mass-change estimates gives a total mass loss for the CAA of $368 \pm 41 \text{ Gt}$ or $1.01 \pm 0.11 \text{ mm}$ sea-level rise for the years 2004 to 2009. Most of the mass loss came from the northern CAA, which lost $224 \pm 30 \text{ Gt}$, with the remaining $144 \pm 28 \text{ Gt}$ coming from the southern CAA (see Supplementary Figs 1–3 for a further subdivision of the mass losses within the northern and southern CAA). We estimate that the majority of the mass loss (about 92%) is due to meltwater runoff, with a much smaller contribution coming from ice discharge from marine-terminating glaciers (about 8%). Three-quarters of all mass loss occurred in the last three years of the observation period with an average loss of $92 \pm 12 \text{ Gt yr}^{-1}$, or $0.25 \pm 0.03 \text{ mm yr}^{-1}$ sea-level rise. This rate is four times greater than the estimated mass loss for CAA over the period 1995 to 2000 (ref. 9).

This increase in mass loss is in direct response to warmer surface air temperatures in summer, to which the glaciers of the CAA have a high sensitivity. Over the six-year period of our study an additional $64 \pm 14 \text{ Gt yr}^{-1}$ of ice was lost to the oceans for every 1 K rise in mean summer surface air temperature. Dividing by the total glacier area gives an area-averaged temperature sensitivity of $-430 \pm 90 \text{ kg m}^{-2} \text{ yr}^{-1} \text{ K}^{-1}$, which is two times larger than estimated from glacier surface mass-budget records^{2,24,25} and is close to sensitivities estimated from regional climatology². The sensitivity to precipitation is much smaller; a 10% increase in precipitation would result in a mass gain of only about 5 Gt yr^{-1} . Such a low sensitivity to precipitation is in contrast to glaciers located in wet maritime regions. For example a 10% increase in precipitation over the Patagonia icefields, which have a combined ice area that is one-tenth the size of the CAA, would result in a 12 Gt yr^{-1} gain of mass²⁶.

To put the mass losses occurring in the CAA into a global perspective, the Patagonia icefields lost ice at an average rate of $28 \pm 11 \text{ Gt yr}^{-1}$ between April 2002 and December 2006 (ref. 27) with little change in the ice-loss trend for the years 2007 to 2009 (J. Chen, personal communication). The glaciers of the Gulf of Alaska lost mass at an average rate of $88 \pm 15 \text{ Gt yr}^{-1}$ for the years 2004 to 2006, slowing to $70 \pm 11 \text{ Gt yr}^{-1}$ for the years 2007 to 2009 (update to ref. 28). The sharp increase in mass loss from the CAA and the slowdown in loss from the Gulf of Alaska makes the CAA the largest contributor to eustatic sea level rise outside Greenland and Antarctica for the years 2007–2009. Because of the high sensitivity to temperature and low sensitivity to precipitation, the CAA is expected to continue to be one of the largest contributing regions to eustatic sea level rise well into the next century and beyond³.

METHODS SUMMARY

The surface mass-budget model was run at a resolution of 500 m by 500 m for the period 1949 to 2009 (Supplementary Information). Model results are validated against observations and agree well with *in situ* point surface mass-budget measurements (Supplementary Fig. 4: $r = 0.86$, $N = 3,717$, standard error = 350 kg m^{-2}). For the four regions with well-established surface mass-budget measurement programmes (Agassiz Ice Cap, north-western Devon Ice Cap, Meighen Ice Cap and White Glacier^{7,8}) the model has a very low bias ($-18 \text{ kg m}^{-2} \text{ yr}^{-1}$) in the glacier-averaged surface mass budget (Supplementary Information). To be consistent with

the other data sets presented in this study, we discuss only mass changes modelled over the ICESat and GRACE operational period between autumn 2003 and autumn 2009.

To recover mass changes from the GRACE measurements we use forward modelling of mass changes in predefined basins, minimizing the least-squares difference between GRACE observations and the forward model in an iterative method (Supplementary Information and refs 29 and 30). To avoid biases from surrounding areas (Supplementary Fig. 1) as a result of the limited spatial resolution and integral character of the GRACE observations, mass changes are modelled for the Greenland Ice Sheet and other areas surrounding the CAA. GRACE measurements were made available by the Center for Space Research (CSR version RL04) and were downloaded from http://podaac.jpl.nasa.gov/DATA_CATALOG/graceinfo.html.

More details about the data and methods can be found in the Supplementary Information.

Received 23 November 2010; accepted 4 April 2011.

Published online 20 April 2011.

- Meier, M. F. *et al.* Glaciers dominate eustatic sea-level rise in the 21st century. *Science* **317**, 1064–1067 (2007).
- Hock, R., de Woul, M., Radić, V. & Dyurgerov, M. Mountain glaciers and ice caps around Antarctica make a large sea level rise contribution. *Geophys. Res. Lett.* **36**, L07501 (2009).
- Kaser, G., Cogley, J. G., Dyurgerov, M. B., Meier, M. F. & Ohmura, A. Mass balance of glaciers and ice caps: consensus estimates for 1961–2004. *Geophys. Res. Lett.* **33**, L19501 (2006).
- Bahr, D. B., Dyurgerov, M. & Meier, M. F. Sea-level rise from glaciers and ice caps: a lower bound. *Geophys. Res. Lett.* **36**, L03501 (2009).
- Radić, V. & Hock, R. Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise. *Nature Geosci.* **4**, 91–94 (2011).
- Radić, V. & Hock, R. Regional and global volumes of glaciers derived from statistical upscaling of glacier inventory data. *J. Geophys. Res.* **115**, doi:10.1029/2009JF001373 (2010).
- Koerner, R. M. Mass balance of glaciers in the Queen Elizabeth Islands, Nunavut, Canada. *Ann. Glaciol.* **42**, 417–423 (2005).
- Cogley, J. G., Adams, W. P., Ecclestone, M. A., Jung-Rothenhäusler, F. & Ommanney, C. S. L. Mass balance of White Glacier, Axel Heiberg Island, NWT, Canada, 1960–91. *J. Glaciol.* **42**, 548–563 (1996).
- Abdalati, W. *et al.* Elevation changes of ice caps in the Canadian Arctic Archipelago. *J. Geophys. Res.* **109**, F04007 (2004).
- Cogley, J. G. *et al.* *Glossary of Glacier Mass Balance and Related Terms*. IHP-VII Technical Documents in Hydrology No. 86 (IACS Contribution No. 2, UNESCO-IHP, in the press).
- Zwally, H. J. *et al.* ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *J. Geodyn.* **34**, 405–445 (2002).
- Lotz, J. R. & Sagar, R. B. Northern Ellesmere Island: an Arctic desert. *Geogr. Ann.* **44**, 366–377 (1962).
- Bradley, R. S. & England, J. Recent climatic fluctuations of the Canadian High Arctic and their significance for glaciology. *Arct. Alp. Res.* **10**, 715–731 (1978).
- Hooke, R. L., Johnson, G. W., Brugger, K. A., Hanson, B. & Holdsworth, G. Changes in mass balance, velocity, and surface profile along a flow line on Barnes Ice Cap, 1970–1984. *Can. J. Earth Sci.* **24**, 1550–1561 (1987).
- Gardner, A. S. & Sharp, M. Influence of the Arctic Circumpolar Vortex on the mass balance of Canadian High Arctic glaciers. *J. Clim.* **20**, 4586–4598 (2007).
- Taylor Alt, B. Developing synoptic analogs for extreme mass balance conditions on Queen Elizabeth Island ice caps. *J. Clim. Appl. Meteorol.* **26**, 1605–1623 (1987).
- Hock, R. Temperature index melt modelling in mountain areas. *J. Hydrol.* **282**, 104–115 (2003).
- Braithwaite, R. J. Positive degree-day factors for ablation on the Greenland Ice Sheet studied by energy-balance modeling. *J. Glaciol.* **41**, 153–160 (1995).
- Gardner, A. S. *et al.* Near-surface temperature lapse rates over Arctic glaciers and their implications for temperature downscaling. *J. Clim.* **22**, 4281–4298 (2009).
- Zwally, H. J. *et al.* GLAS/ICESat L1B Global Elevation Data V031, 20 February 2003 to 11 October 2009 (National Snow and Ice Data Center, 2010).
- Moholdt, G., Nuth, C., Hagen, J. O. & Kohler, J. Recent elevation changes of Svalbard glaciers derived from ICESat laser altimetry. *Remote Sens. Environ.* **114**, 2756–2767 (2010).
- Arendt, A. *et al.* Updated estimates of glacier volume changes in the western Chugach Mountains, Alaska, and a comparison of regional extrapolation methods. *J. Geophys. Res.* **111**, F000436 (2006).
- Sole, A., Payne, T., Bamber, J., Nienow, P. & Krabill, W. Testing hypotheses of the cause of peripheral thinning of the Greenland Ice Sheet: is land-terminating ice thinning at anomalously high rates? *Cryosphere* **2**, 205–218 (2008).
- Oerlemans, J. *et al.* Estimating the contribution of Arctic glaciers to sea-level change in the next 100 years. *Ann. Glaciol.* **42**, 230–236 (2005).
- De Woul, M. & Hock, R. Static mass-balance sensitivity of Arctic glaciers and ice caps using a degree-day approach. *Ann. Glaciol.* **42**, 217–224 (2005).
- Rignot, E., Rivera, A. & Casassa, G. Contribution of the Patagonia icefields of South America to sea level rise. *Science* **302**, 434–437 (2003).
- Chen, J. L., Wilson, C. R., Tapley, B. D., Blankenship, D. D. & Ivins, E. R. Patagonia icefield melting observed by gravity recovery and climate experiment (GRACE). *Geophys. Res. Lett.* **34**, L22501 (2007).

28. Luthcke, S. B., Arendt, A. A., Rowlands, D. D., McCarthy, J. J. & Larsen, C. F. Recent glacier mass changes in the Gulf of Alaska region from GRACE mascon solutions. *J. Glaciol.* **54**, 767–777 (2008).
29. Wouters, B., Chambers, D. & Schrama, E. J. O. GRACE observes small-scale mass loss in Greenland. *Geophys. Res. Lett.* **35**, L20501 (2008).
30. van den Broeke, M. *et al.* Partitioning recent Greenland mass loss. *Science* **326**, 984–986 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Arendt for reviewing the manuscript and S. Luthcke and A. Arendt for providing the updated glacier mass anomalies for Alaska. We thank H. Blatter, W. Colgan, E. Dowdeswell, M. Huss, S. Marshall and D. Mueller for contributing observational data sets. We thank R. Riva and P. Stocchi for providing glacial isostatic adjustment models. This work was supported by funding to A.S.G. from NSERC Canada and the Alberta Ingenuity Fund, funding to

G.M. by the European Union 7th Framework Program (grant number 226375) through the ice2sea programme (contribution number 017), and funding to M.J.S. from NSERC and CFCAS (through the Polar Climate Stability Network). The SMB modelling was conducted using the infrastructure and resources of AICT of the University of Alberta.

Author Contributions A.S.G. developed the study and wrote the paper. A.S.G., G.M. and B.W. all contributed equally to the analysis, using SMB+D, ICESat and GRACE, respectively. G.J.W. provided ice and basin outlines, model topography and created Fig. 1. The remaining authors provided *in situ* measurements. All authors discussed and commented on the manuscript at all stages.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.S.G. (alexsg@umich.edu).

Melting of the Earth's inner core

David Gubbins^{1,2}, Binod Sreenivasan³, Jon Mound¹ & Sebastian Rost¹

The Earth's magnetic field is generated by a dynamo in the liquid iron core, which convects in response to cooling of the overlying rocky mantle. The core freezes from the innermost surface outward, growing the solid inner core and releasing light elements that drive compositional convection^{1–3}. Mantle convection extracts heat from the core at a rate that has enormous lateral variations⁴. Here we use geodynamo simulations to show that these variations are transferred to the inner-core boundary and can be large enough to cause heat to flow into the inner core. If this were to occur in the Earth, it would cause localized melting. Melting releases heavy liquid that could form the variable-composition layer suggested by an anomaly in seismic velocity in the 150 kilometres immediately above the inner-core boundary^{5–7}. This provides a very simple explanation of the existence of this layer, which otherwise requires additional assumptions such as locking of the inner core to the mantle, translation from its geopotential centre^{7,8} or convection with temperature equal to the solidus but with composition varying from the outer to the inner core⁹. The predominantly narrow downwellings associated with freezing and broad upwellings associated with melting mean that the area of melting could be quite large despite the average dominance of freezing necessary to keep the dynamo going. Localized melting and freezing also provides a strong mechanism for creating seismic anomalies in the inner core itself, much stronger than the effects of variations in heat flow so far considered¹⁰.

The core responds passively to the non-uniform heat flow imposed by the mantle: it plays a purely passive role in this coupled convective system. Variations in heat flux around the core–mantle boundary (CMB), created by mantle convection, are likely to be large. They can be estimated by two independent methods, one using seismic tomography¹¹ within the supposed thermal boundary layer at the base of the mantle, and the other using mantle convection studies⁴. Both suggest variations comparable with the average heat flux. Inhomogeneous boundary conditions can produce enormous effects on core convection^{12–14}, and when background convection is small the boundary variations can aid magnetic field generation through enhanced helical motions in fluid columns¹⁵. Many geodynamo simulations have incorporated thermal boundary conditions based on seismic tomography to explain the non-axisymmetric time average of the geomagnetic field^{16–18}, low secular variation in the Pacific^{16,19}, frequency of polarity reversals²⁰, and persistent polarity transition paths during reversals²¹.

We have explored the heat flux variability on the inner-core boundary (ICB) using numerical geodynamo calculations driven by thermal convection with an inhomogeneous upper boundary heat flux and constant lower boundary temperature. The details of our dynamo model are given in the Methods. Examples using the ‘tomographic’ boundary condition¹¹ suffice to illustrate the possibility of inward heat flow at the lower boundary. The important parameter $q^* = (q_{\max} - q_{\min})/2q_{\text{mean}}$ measures the strength of the lateral variation in CMB heat flux relative to the average; a range from $q^* = 0.15$ to 0.45 gives dynamos that vary from one relatively unaffected by the boundary condition to one where the magnetic field is almost stationary, or statistically ‘locked’ to the boundary²².

Figure 1 gives the heat flux distribution on the upper and lower boundaries for a locked dynamo at $q^* = 0.45$. The pattern of heat flux on the ICB mirrors that on the CMB; negative patches of heat flux indicate heat flow into the inner core at sites of melting if this were part of the model. Figure 2 shows two snapshots and a time average for a dynamo with $q^* = 0.15$; again there are patches where the heat flux is negative despite the weaker lateral variations. Upwellings in the outer core are broad while downwellings are narrow and vertical in all these dynamos (Fig. 3), producing concentrated patches of high ICB heat flux immediately beneath high CMB heat flux. The regions of melting are therefore relatively large in comparison with the total amount of melting. We note, however, that dynamo models with different operating parameters and buoyancy profiles need not produce heat flowing into the lower boundary: a weakly convecting regime in which lateral variations at the upper boundary are allowed to propagate all the way to the lower boundary appears to be the most conducive for inner-core melting.

Three complications must be taken into account when applying the results of a thermal geodynamo simulation to the Earth. The first is the

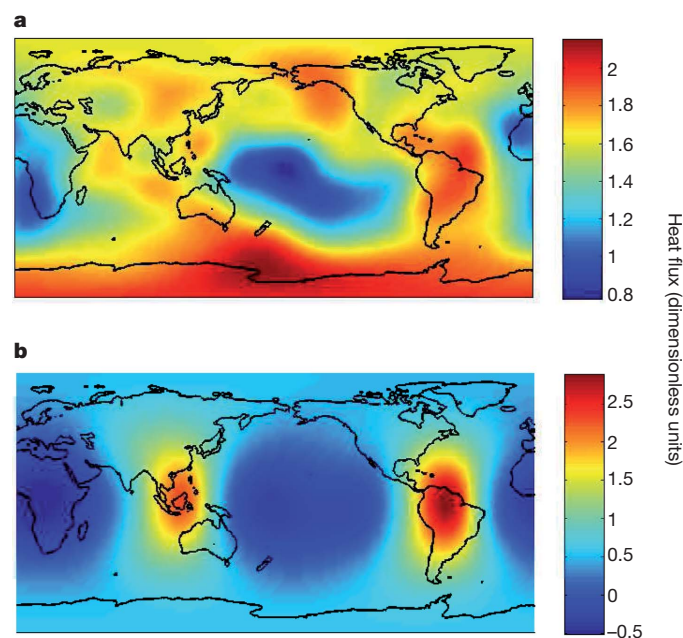


Figure 1 | Effect of mantle inhomogeneity on heat flux distribution at the inner core surface. Heat fluxes are applied to the upper boundary (a) and calculated on the constant-temperature lower boundary (b) in a geodynamo simulation where the flow is strongly coupled to the boundary thermal anomalies ($q^* = 0.45$). The range of heat flux across the upper boundary ranges from 0.77 to 2.16 dimensionless units outwards and across the lower boundary ranges from -0.51 to 2.89 dimensionless units (negative values indicate heat flux into the inner core). This model uses an Ekman number 1.2×10^{-4} , Rayleigh number 1.5 times the critical value for onset of convection, Prandtl number 1 and magnetic Prandtl number 10. (See the Methods section for definitions of these dimensionless numbers.)

¹School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK. ²IGPP, SIO, UCSD 9500 Gilman Drive, La Jolla, California 92093-0235, USA. ³Department of Mechanical Engineering, Indian Institute of Technology, Kanpur 208 016, India.

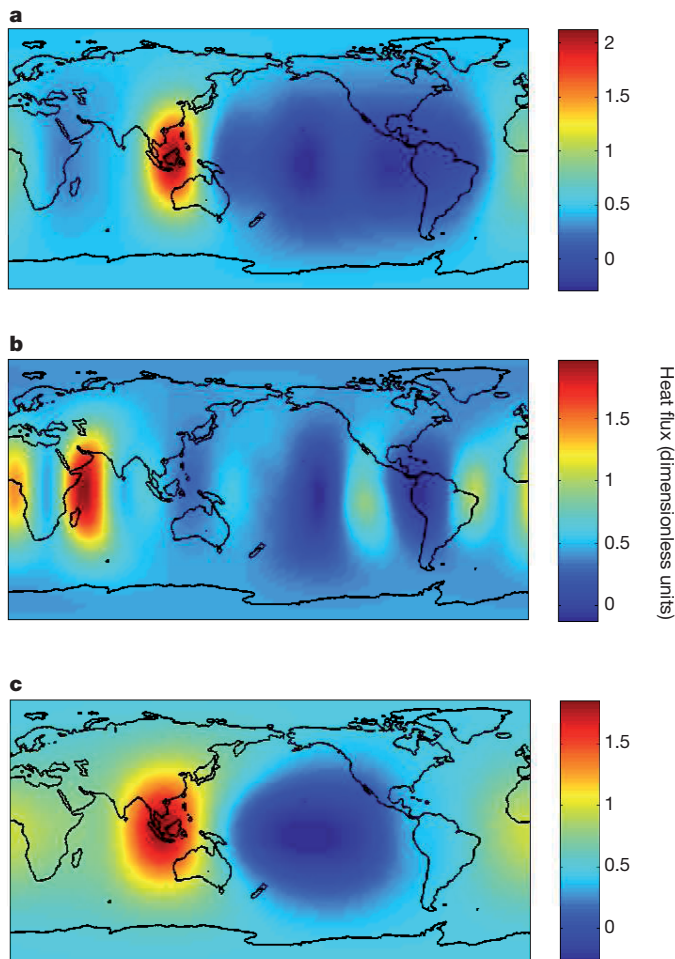


Figure 2 | Calculated heat flux on the lower boundary of a geodynamo model where $q^* = 0.15$ for the upper boundary heat flux. Panels **a** and **b** are snapshots and **c** shows the time average over several magnetic diffusion times. Heat fluxes range from -0.287 to 2.126 (**a**), -0.124 to 1.976 (**b**) and -0.276 to 1.86 (**c**) for the time average. The parameters used in this model are the same as in Fig. 1.

heat conducted down the adiabat. This was omitted from a recent mantle convection study that explored the effects of a postperovskite layer and variations in chemical composition on the heat flux across the CMB and its correlation with seismic shear wave velocity⁴. Postperovskite makes little difference to heat fluxes but lateral variations in composition, such as a subducted slab lying on the CMB, greatly increase the ratio q^* . To apply these results to core convection we must first subtract the heat conducted down the adiabatic temperature gradient. Typical estimates of the adiabatic gradient at the CMB (1 K km^{-1}) and core thermal conductivity ($k = 50 \text{ W m}^{-1} \text{ K}^{-1}$) give a conducted heat flux of 50 mW m^{-2} , comparable with q_{mean} for the mantle convection calculations. Subtracting this raises the relevant q^* dramatically because it reduces q_{mean} to nearly zero while leaving the range $q_{\text{max}} - q_{\text{min}}$ unchanged. In fact there is nothing to stop q^* becoming infinite, as it nearly does for the most realistic mantle model in the previous study⁴ (model TC-3.6, which has a compressible pyroxene content), it merely means the top of the core is thermally neutral. Most dynamo simulations have been restricted to rather low q^* because the dynamo tends to fail for large lateral heat flux variations^{15,18}. In our models with internal heating the dynamo fails by $q^* \approx 1$ but dynamos with basal heating and stratified upper layers continue to work for large q^* (ref. 23). The upper region of the Earth's fluid core is likely to be stably stratified, or at most only weakly convecting^{24,25}, and a high q^* is therefore quite possible and appropriate for the Earth. Two factors are likely to increase q^* with depth. First,

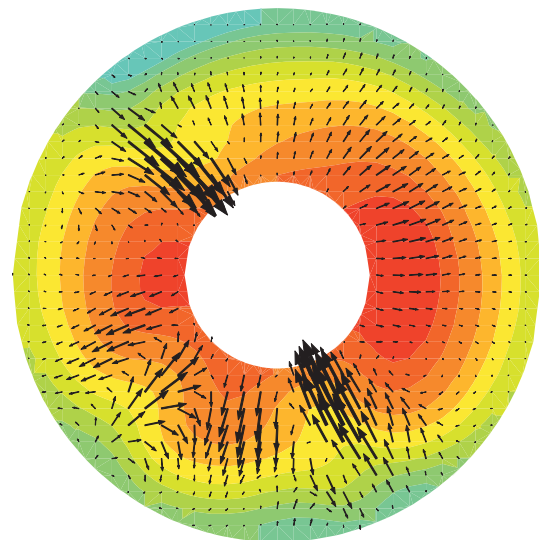


Figure 3 | Temperature (colour contours) and fluid flow (arrows) on the equatorial section for the statistically locked tomographic model ($q^* = 0.45$). The lowest temperature is blue and the highest temperature is deep red. We note the narrow downwellings beneath cold regions (the two major ones coinciding with the 'ring of fire' around the Pacific) and broad upwellings (corresponding to the mid-Pacific and African superplume). This leads to relatively large areas of negative (melting) and low-positive heat flux on the ICB and relatively small areas of strong-positive heat flux (freezing).

the adiabatic gradient weakens with depth by a factor of about three between the CMB and ICB. At the ICB the adiabatic heat flux must be added back on to the model results, reducing any heat flow into the inner core; however, the weakened adiabat makes this a relatively small effect. Second, narrow downwellings and the spherical geometry tend to concentrate the convected heat flux, increasing the lateral variations.

The second complication to consider is compositional convection. The compositional gradient is neutral or stabilizing at the CMB (assuming no passage of light elements into the mantle): convection at the top of the core is purely thermal. Compositional buoyancy tends to dominate thermal buoyancy deeper in the outer core, particularly near the ICB, as the following calculation shows. The buoyancy force is $\rho(\alpha_c c + \alpha_T T)g$, where ρ is the density, g the acceleration due to gravity, α_T the thermal expansion coefficient and α_c the compositional expansion coefficient. Compositional changes therefore have the thermal equivalent $\alpha_c c / \alpha_T$. Comparing heat and mass fluxes in the respective diffusion equations show that the conversion factor is $C_p \alpha_c / \alpha_T$, where C_p is the specific heat. Freezing 1 kg of liquid at the ICB releases L joules of latent heat and ρc kilograms of mass with thermal equivalent $C_p \alpha_c c / \alpha_T$ joules. The effective buoyancy ratio is therefore $C_p \alpha_c c / L \alpha_T = 2.3$ for a concentration $c = 0.0252$, corresponding to a density jump at the ICB of 0.6 g cm^{-3} (from PREM²⁶) assuming that 0.34 g cm^{-3} of this comes from the solid-liquid phase transition for pure iron²⁷. Compositional buoyancy dominates and will be even larger for larger ICB density jumps: 4.1 for 0.8 g cm^{-3} and 5.8 for 1.0 g cm^{-3} . Thus temperature variations are relatively unimportant in the buoyancy force near the ICB but are crucial in determining the rate of freezing, and therefore the supply of buoyancy through the release of light elements. Lateral variations in temperature imposed by the upper boundary will be carried down to the ICB by compositional convection, assisted by thermal convection, so we expect the variations on the ICB observed in the thermal or codensity geodynamo simulations to be sustained in a thermo-chemical system.

The third complication is the possible dynamic consequences of the variable-composition layer. The density gradient across the layer of freshly melted, heavy liquid is vastly steeper than anything arising from convection in the main part of the outer core: a density change of the order of 0.1 g cm^{-3} across a 150-km layer compared to a typical

convective density fluctuation of $10^{-6} \text{ g cm}^{-3}$ or less across a comparable or longer length scale, as estimated from a buoyancy–Coriolis force balance near the ICB²⁸. Such a steep density gradient would prevent downwellings from reaching the ICB, but plumes of light material produced by freezing would rise through it, drawing heavy liquid along the ICB towards the regions of freezing and maintaining mixing of the variable-composition layer. Laboratory experiments suggest that the plumes could mix the layer if the melting exceeds 20% of the freezing⁸, but the plumes on the ICB are determined by thermal, not compositional, effects. Further study is needed to understand the influence of this layer.

Regional melting of the inner core that results from heat flux variations at the CMB provides the simplest explanation of the observed variable-composition layer at the base of the outer core. It also provides a strong mechanism for seismic anomalies in the solid inner core itself because areas of melting will consist of recently exposed, precompressed material whereas areas of freezing will have layers of recently formed, unconsolidated mush. Variations in heat flux have already been invoked to explain seismic anomalies inside the inner core²⁹ but actual melting will produce even stronger effects⁷. In both cases, any correlation with mantle anomalies and persistence of locality requires the inner core and, to some extent, the core flow to be locked to the mantle. If these observations hold up to further scrutiny—in particular, if the variable-composition layer turns out not to require inner-core locking—they will provide important constraints on core evolution, convection and the dynamo.

METHODS SUMMARY

We consider a thermal convection-driven dynamo operating in an electrically conducting fluid. The Earth's outer core is modelled as a spherical shell confined between a solid iron inner core of radius r_i and an insulating mantle at radius r_o . The radius ratio r_i/r_o is taken to be that of the Earth, 0.35. In the Boussinesq approximation³⁰, the time-dependent, three-dimensional magnetohydrodynamic equations for the velocity \mathbf{u} , the magnetic field \mathbf{B} and the temperature T are solved numerically. The governing equations and numerical method are described in the Methods. The inner boundary in the model is considered to be at a fixed temperature, whereas the outer boundary is subject to a lateral variation in heat flux that has the same structure as the seismic shear-wave velocity variation in the lower mantle¹¹. This assumes shear velocity is determined by temperature and not by composition. The dominant pattern is a fast (cold) ring around the Pacific rim and slow (hot) regions beneath the Pacific and Africa (see Fig. 1).

The parameter regime used in this paper has been considered in the study of a boundary-locked dynamo^{15,22}. The Ekman number is kept sufficiently small to make the dynamics rotationally dominant, and the Rayleigh number is chosen such that free convection does not swamp the effect of the CMB lateral inhomogeneity. When the heat flux inhomogeneity ratio q^* is sufficiently large, this regime is characterized by a boundary-driven thermal wind balance¹⁵, that is, a balance between the lateral buoyancy and Coriolis forces. This force balance causes the narrow downwellings to remain locked at preferred longitudes, which can explain the quasi-stationary, non-axisymmetric flux patches in today's geomagnetic field.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 February; accepted 23 March 2011.

1. Braginsky, S. I. Structure of the F layer and reasons for convection in the Earth's core. *Dokl. Akad. Nauk. SSSR (Engl. Trans.)* **149**, 1311–1314 (1963).
2. Labrosse, S., Poirier, J.-P. & Le Mouél, J.-L. On cooling of the Earth's core. *Phys. Earth Planet. Inter.* **99**, 1–17 (1997).
3. Nimmo, F., Price, G. D., Brodholt, J. & Gubbins, D. The influence of potassium on core and geodynamo evolution. *Geophys. J. Int.* **156**, 363–376 (2004).

4. Nakagawa, T. & Tackley, P. J. Lateral variations in CMB heat flux and deep mantle seismic velocity caused by a thermal-chemical-phase boundary layer in 3D spherical convection. *Earth Planet. Sci. Lett.* **271**, 348–358 (2008).
5. Souriau, A. & Poupinet, G. The velocity profile at the base of the liquid core from PKP(BC+Cdiff) data: an argument in favor of radial inhomogeneity. *Geophys. Res. Lett.* **18**, 2023–2026 (1991).
6. Song, X. D. & Helmberger, D. V. Velocity structure near the inner core boundary from waveform modeling. *J. Geophys. Res.* **97**, 6573–6586 (1992).
7. Monnereau, M., Calvet, M., Margerin, L. & Souriau, A. Lopsided growth of Earth's inner core. *Science* **328**, 1014–1017 (2010).
8. Alboussière, T., Deguen, R. & Melzani, M. Melting-induced stratification above the Earth's inner core due to convective translation. *Nature* **466**, 744–747 (2010).
9. Gubbins, D., Masters, G. & Nimmo, F. A thermochemical boundary layer at the base of Earth's outer core and independent estimate of core heat flux. *Geophys. J. Int.* **174**, 1007–1018 (2008).
10. Aubert, J., Amit, H. & Hulot, G. Detecting thermal boundary control in surface flows from numerical dynamos. *Phys. Earth Planet. Inter.* **160**, 143–156 (2007).
11. Masters, T. G., Johnson, S., Laske, G. & Bolton, H. F. A shear-velocity model of the mantle. *Phil. Trans. R. Soc. Lond. A* **354**, 1385–1411 (1996).
12. Zhang, K. & Gubbins, D. Convection in a rotating spherical fluid shell with an inhomogeneous temperature boundary condition at finite Prandtl number. *Phys. Fluids* **8**, 1141–1148 (1996).
13. Zhang, K. & Gubbins, D. Scale disparities and magnetohydrodynamics in the Earth's core. *Phil. Trans. R. Soc. Lond. A* **358**, 899–920 (2000).
14. Gibbons, S. & Gubbins, D. Convection in the Earth's core driven by lateral variations in the core-mantle boundary heat flux. *Geophys. J. Int.* **142**, 631–642 (2000).
15. Sreenivasan, B. On dynamo action produced by boundary thermal coupling. *Phys. Earth Planet. Inter.* **177**, 130–138 (2009).
16. Bloxham, J. The effect of thermal core-mantle interactions on the paleomagnetic secular variation. *Phil. Trans. R. Soc. Lond. A* **358**, 1171–1179 (2000).
17. Christensen, U., Olson, P. & Glatzmaier, G. A. A dynamo model interpretation of geomagnetic field structures. *Geophys. Res. Lett.* **25**, 1565–1568 (1998).
18. Olson, P. & Christensen, U. R. The time-averaged magnetic field in numerical dynamos with non-uniform boundary heat flow. *Geophys. J. Int.* **151**, 809–823 (2002).
19. Christensen, U. R. & Olson, P. Secular variation in numerical geodynamo models with lateral variations of boundary heat flow. *Phys. Earth Planet. Inter.* **138**, 39–54 (2003).
20. Glatzmaier, G. A., Coe, R. S., Hongre, L. & Roberts, P. H. The role of the Earth's mantle in controlling the frequency of geomagnetic reversals. *Nature* **401**, 885–890 (1999).
21. Kutzner, C. & Christensen, U. R. Simulated geomagnetic reversals and preferred virtual geomagnetic pole paths. *Geophys. J. Int.* **157**, 1105–1118 (2004).
22. Gubbins, D., Willis, A. P. & Sreenivasan, B. Correlation of Earth's magnetic field with lower mantle thermal and seismic structure. *Phys. Earth Planet. Inter.* **162**, 256–260 (2007).
23. Sreenivasan, B. & Gubbins, D. Dynamos with weakly convecting outer layers: implications for core–mantle boundary interaction. *Geophys. Astrophys. Fluid Dyn.* **102**, 395–407 (2008).
24. Buffett, B. A. & Seagle, C. T. Stratification of the top of the core due to chemical interactions with the mantle. *J. Geophys. Res.* **115**, B04407 (2010).
25. Helffrich, G. & Kaneshima, S. Outer-core compositional stratification from observed core wave speed profiles. *Nature* **468**, 807–810 (2010).
26. Dziewonski, A. M. & Anderson, D. L. Preliminary Reference Earth Model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
27. Gubbins, D., Alfè, D., Masters, T. G. & Price, D. Gross thermodynamics of 2-component core convection. *Geophys. J. Int.* **157**, 1407–1414 (2004).
28. Moffatt, H. K. & Loper, D. E. The magnetostrophic rise of a buoyant parcel in the Earth's core. *Geophys. J. Int.* **117**, 394–402 (1994).
29. Aubert, J., Amit, H., Hulot, G. & Olson, P. Thermochemical flows couple the Earth's inner core growth to mantle heterogeneity. *Nature* **454**, 758–761 (2008).
30. Sreenivasan, B. A buoyant flow structure in a magnetic field: quasi-steady states and linear–nonlinear transitions. *Phys. Lett. A* **372**, 5471–5478 (2008).

Author Contributions B.S. set up the geodynamo model and performed the calculations that form the basis of this paper. All four authors discussed the results and contributed to the text of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.S. (bsreeni@iitk.ac.in).

METHODS

Numerical dynamo model. We consider a thermal convection-driven dynamo in which an electrically conducting fluid is confined between two concentric, co-rotating spherical surfaces. The radius ratio r_i/r_o is chosen to be that of the Earth, 0.35. In the Boussinesq approximation³⁰, the time-dependent, three-dimensional magnetohydrodynamic equations for the velocity \mathbf{u} , the magnetic field \mathbf{B} and the temperature T are solved numerically³¹. The governing dimensionless equations are:

$$\frac{E}{\text{Pm}} \left(\frac{\partial \mathbf{u}}{\partial t} + (\nabla \times \mathbf{u}) \times \mathbf{u} \right) + 2\hat{\mathbf{z}} \times \mathbf{u} = -\nabla p + \text{Ra Pm Pr}^{-1} T \mathbf{r} + (\nabla \times \mathbf{B}) \times \mathbf{B} + E \nabla^2 \mathbf{u} \quad (1)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \nabla^2 \mathbf{B} \quad (2)$$

$$\frac{\partial T}{\partial t} + (\mathbf{u} \cdot \nabla) T = \text{Pm Pr}^{-1} \nabla^2 T \quad (3)$$

$$\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{B} = 0 \quad (4)$$

The dimensionless groups in the above equations are the Ekman number, $E = \nu/2\Omega D^2$, the Prandtl number, $\text{Pr} = \nu/\kappa$, the magnetic Prandtl number, $\text{Pm} = \nu/\eta$ and the ‘modified’ Rayleigh number $\text{Ra} = g\alpha\beta_i D^3/2\Omega\kappa$, which is the product of the conventional Rayleigh number and the Ekman number. The definition of the Rayleigh number depends on the basic state (conductive) profile in the model (see below). In the above dimensionless groups, ν is the kinematic viscosity, κ is the thermal diffusivity, η is the magnetic diffusivity, D is the gap-width of the spherical shell, Ω is the angular velocity of rotation, g is the gravitational acceleration, α is the coefficient of thermal expansion and β_i is a constant that determines the basic state temperature profile, T_0 . The Ekman number is a

measure of the rotation rate and the Rayleigh number represents the strength of convective buoyancy in the problem. Our models use an Ekman number of $E = 1.2 \times 10^{-4}$, a Rayleigh number of 1.5Ra_c , where Ra_c is the critical Rayleigh number for onset of nonmagnetic convection, a Prandtl number $\text{Pr} = 1$ and magnetic Prandtl number $\text{Pm} = 10$.

No-slip boundary conditions are imposed on the flow at the ICB and at the CMB. The inner core is considered to be at a fixed temperature and electrically conducting. The isothermal condition at the ICB is reasonable for a solid core of high thermal conductivity. However, compositional buoyancy in the form of light-element release over areas of freezing can complicate the boundary condition at the ICB. The upper boundary in the model is maintained electrically insulating to mimic the mantle and subject to a lateral variation in heat flux that has the same structure as the seismic shear-wave velocity variation in the lower mantle. The basic state temperature profile imposed in the model represents a uniform distribution of heat sources, and is given by $T_0(r) = \beta_i(r_i^2 - r^2)/2$, where r_i is the inner radius and β_i is related to a prescribed, uniform heat source Q_s as follows: $\beta_i = Q_s/3\kappa$.

The velocity and magnetic field vectors are expressed in terms of poloidal and toroidal scalars, as follows:

$$\begin{aligned} \mathbf{u} &= \nabla \times \nabla \times [P_u \mathbf{r}] + \nabla \times [T_u \mathbf{r}] \\ \mathbf{B} &= \nabla \times \nabla \times [P_B \mathbf{r}] + \nabla \times [T_B \mathbf{r}] \end{aligned} \quad (5)$$

whereby the continuity equations (4) are satisfied. The standard numerical method used here involves expanding the above four scalar variables and the temperature T in spherical harmonics in latitude θ and longitude ϕ , and time-stepping the spectral coefficients. Finite differences are used in the radial direction. The numerical integration of the equations is performed for at least five magnetic diffusion times.

31. Sreenivasan, B. & Jones, C. A. The role of inertia in the evolution of spherical dynamos. *Geophys. J. Int.* **164**, 467–476 (2006).

Eocene lizard from Germany reveals amphisbaenian origins

Johannes Müller¹, Christy A. Hipsley^{1,2}, Jason J. Head³, Nikolay Kardjilov⁴, André Hilger⁴, Michael Wuttke⁵ & Robert R. Reisz³

Amphisbaenia is a speciose clade of fossorial lizards characterized by a snake-like body and a strongly reinforced skull adapted for head-first burrowing^{1,2}. The evolutionary origins of amphisbaenians are controversial, with molecular data uniting them with lacertids^{3,4}, a clade of Old World terrestrial lizards, whereas morphology supports a grouping with snakes and other limbless squamates^{5–9}. Reports of fossil stem amphisbaenians¹⁰ have been falsified¹¹, and no fossils have previously tested these competing phylogenetic hypotheses or shed light on ancestral amphisbaenian ecology. Here we report the discovery of a new lacertid-like lizard from the Eocene Messel locality of Germany that provides the first morphological evidence for lacertid–amphisbaenian monophyly on the basis of a reinforced, akinetic skull roof and braincase, supporting the view that body elongation and limblessness in amphisbaenians and snakes evolved independently. Morphometric analysis of body shape and ecology in squamates indicates that the postcranial anatomy of the new taxon is most consistent with opportunistically burrowing habits, which in combination with cranial reinforcement indicates that head-first burrowing evolved before body elongation and may have been a crucial first step in the evolution of amphisbaenian fossoriality.

Reptilia Laurenti, 1768

Squamata Oppel, 1811

Lacertibaenia Vidal and Hedges, 2005

Cryptolacerta hassiaca gen. et sp. nov.

Etymology. *Crypto-*, from the ancient Greek κρυπτός, meaning ‘hidden’ or ‘secret’, referring to the inferred ecology of the animal; *lacerta* (Latin), meaning lizard; *hassiaca* (Latin), female adjective for Hesse, the German province of the Messel locality.

Holotype. SMF ME 2604 (Fig. 1), Forschungsinstitut und Naturmuseum Senckenberg, Frankfurt, Germany.

Locality and horizon. West of Quarry 2, 50 cm above Level β¹². Messel Pit World Heritage Site, Hesse, Germany; Eocene (Lutetian).

Diagnosis. Lacertibaenian squamate with a snout–vent length of approximately 7 cm; skull capsule-like, anteriorly downturned and heavily ossified; transverse nasofrontal suture; small narial openings facing strictly anteriorly owing to a unique dorsolateral covering by the maxilla; small posterodorsal coronoid process of the dentary; 14 dentary, 7 premaxillary and 12 maxillary teeth with the posterior-most maxillary tooth enlarged; 27 presacral vertebrae; manus and pes strongly reduced in size relative to the remaining limb. Shares with amphisbaenians a relatively elongated postorbital skull portion, blunt and rounded snout, sutural contact between prefrontal and postorbitofrontal, contact between prefrontal and jugal, absence of a lacrimal, small jugal with only little angulation, subequal width of the anterior and posterior borders of the frontal, absence of frontal constriction between the orbits, loss of the tympanic crest, neural spines reduced, seven or fewer cervical vertebrae, rod-like clavicles, absence of an

anterior coracoid emargination and interclavicle, fusion of cephalic scales, transversely widened frontal subolfactory processes, thickening of maxilla and frontal, small orbits, a vertical tongue-and-groove articulation between the frontals, and absence of an iliac anterodorsal projection.

The type and only known specimen of *Cryptolacerta hassiaca* is nearly complete, missing only the distal tail (Fig. 1a). Computed tomography (CT) imaging and specimen examination (Figs 1b, c and 2) reveal a mosaic of lacertid and amphisbaenian anatomical characters. The skull is massive and heavily ossified, with an anteroventrally downturned anterior portion (Fig. 1d). Extensive dermal sculpturing covers the skull roof and well preserved scute sulci reflect the presence of large, transversely oriented scales. Both maxilla and frontals display a massive thickening in cross-section coupled with an increase in bone density, obscuring the vascularized internal structure seen in the more posterior cranial elements as well as in lacertids and most other lizards; the same condition in *Cryptolacerta* occurs in amphisbaenians. The external nares are small and anteriorly oriented and are bounded dorsally by a unique anteromedial flange of the maxilla. Small orbits indicate reduced eyes, and the prefrontal and postfrontal have a strong sutural contact similar to fossil amphisbaenians^{13–15}. *Cryptolacerta* has a vertically tall tongue-and-groove interdigitation of the median contact of the paired frontals as in amphisbaenians (Fig. 2a), and the prominent frontal subolfactory processes, although lacking a median contact as in lacertids, are notably widened transversely and form the major part of the posterior wall of the nasal capsule, a feature shared with amphisbaenians (see Supplementary Information). The parietal table is prominent and shows the typical lacertid Y-shaped crest that articulates with the braincase on its ventral surface (Fig. 2b). The large size and ventral extension of the crest indicates close proximity or ossification with the prootic and supraoccipital, resulting in reduced cranial kinesis. The braincase is crushed, and only the parabasisphenoid and slender basiptyergoid articulations are preserved. The middle ear is reduced as evidenced by the absence of a quadrate tympanic crest (Fig. 2c). The dentary has a posterolateral extension covering the anterolateral part of the coronoid (Fig. 1c), as in many amphisbaenians, despite retaining a typically lacertid shape.

Cryptolacerta possesses a distinctive heterodont dentition. The six preserved teeth on the premaxilla are conical and diminutive. The 11 maxillary teeth continuously increase in size posteriorly, with the last tooth being expanded and bulbous in shape and the remaining teeth having bicuspid crowns. The 14 teeth on the dentary also are similar to the maxillary teeth, but lack an enlarged posterior-most tooth.

Postcranially, *Cryptolacerta* possesses 29 procoelous precaudal vertebrae with very low neural spines, including seven cervicals and two sacals. The pectoral girdle consists of recurved clavicles, slender scapulacoracoids and the sternum, whereas the interclavicle is absent (Fig. 2d). The pelvic girdle possesses a well-developed ilium that lacks an anterodorsal process (Fig. 2e). Although not all autopodia are fully

¹Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung an der Humboldt-Universität zu Berlin, Invalidenstr. 43, 10115 Berlin, Germany. ²Department of Ecology and Evolutionary Biology, Earth and Marine Sciences Building, University of California, Santa Cruz, California 95064, USA. ³Department of Biology, University of Toronto, Mississauga, 3359 Mississauga Road North, Mississauga, Ontario L5L 1C6, Canada. ⁴Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin, Germany. ⁵Generaldirektion Kulturelles Erbe RLP, Direktion Landesarchäologie, Referat Erdgeschichte, Große Langgasse 29, 55116 Mainz, Germany.

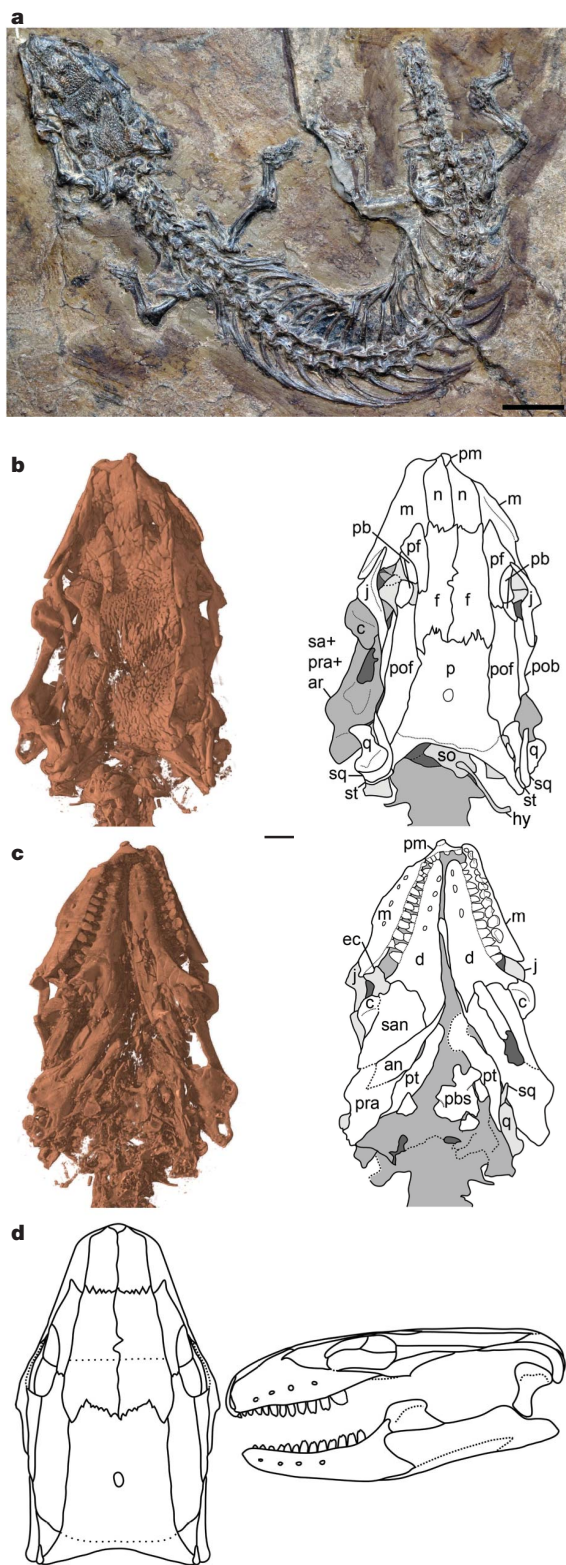


Figure 1 | *Cryptolacerta hassiaca* gen. et sp. nov., holotype (SMF ME 2604). **a**, Nearly complete specimen. **b**, Micro-CT scan of skull and outline of bones in dorsal view. **c**, Micro-CT scan of skull and outline of bones in ventral view. **d**, Reconstruction of skull in dorsal and lateral view. Scale bars, 5 mm (**a**), 2 mm (**b**, **c**). Abbreviations: an, angular; ar, articular; c, coronoid; d, dentary; ec, ectopterygoid; f, frontal; hy, hyoid; j, jugal; m, maxilla; n, nasal; p, parietal; pb, palpebral; pbs, parabasisphenoid; pf, prefrontal; pob, postorbital; pof, postfrontal; pm, premaxilla; pra, prearticular; pt, pterygoid; q, quadrate; sa, surangular; so, supraoccipital; sq, squamosal; st, supratemporal.

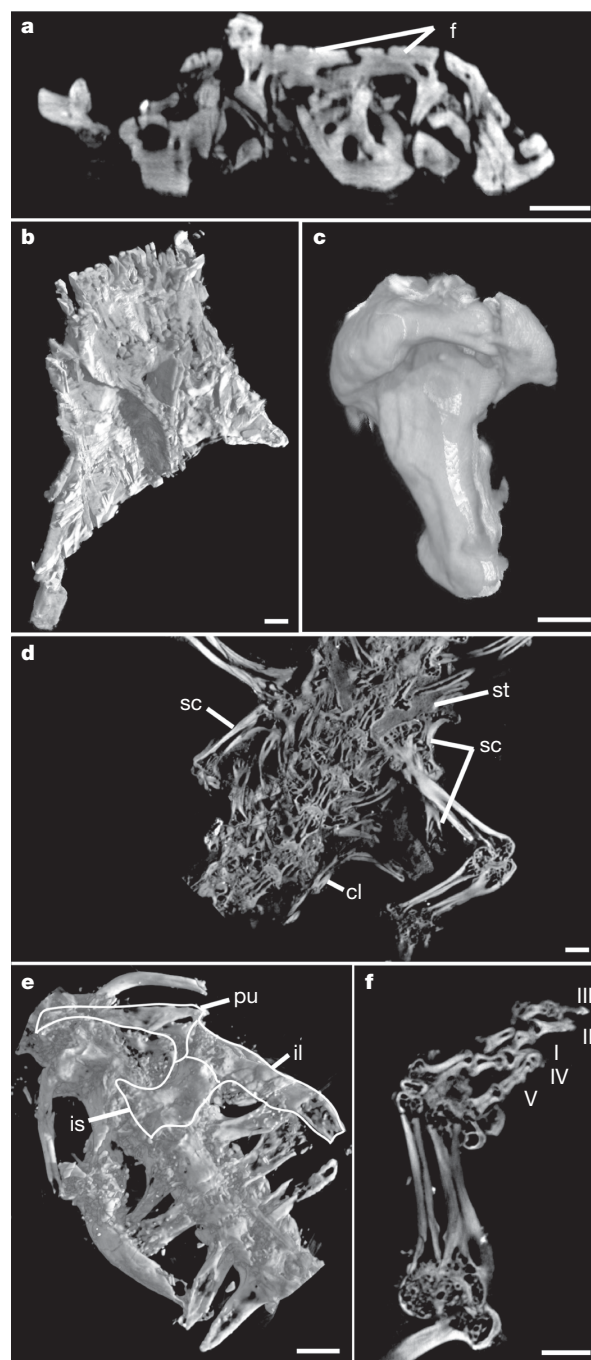


Figure 2 | *Cryptolacerta hassiaca* gen. et sp. nov., holotype (SMF ME 2604), anatomical features as revealed by CT. **a**, Transverse section through the anterior part of the frontals (f). **b**, Parietal in ventral view showing the Y-shaped crest. **c**, Left quadrate in posterolateral view. **d**, Section showing shoulder girdle with sternum (st), scapulocoracoid (sc) and clavicle (cl). **e**, Pelvic girdle with outlines to emphasize the morphology of ischium (is), ilium (il) and pubis (pu). **f**, Manus with digits I–V; note that digit IV lies on top of digit V, as revealed by different CT sections. Scale bars, 1 mm.

preserved, the phalangeal formula of 2–3–4–4/5?–3 suggests that no digits are lost, but the phalangeal elements are miniaturized relative to the remaining limb bones (Fig. 2f).

The systematic position of amphisbaenians within Squamata is poorly constrained. Molecular data support a sister-taxon relationship with lacertids^{3,4,16,17}, but there is no morphological character support among living taxa uniting the highly derived amphisbaenians with lacertids. Most morphological analyses support a common ancestry of amphisbaenians and snakes^{6–9}, but character support for this

hypothesis has been considered homoplastic¹⁸. To determine the significance of *Cryptolacerta* for resolving systematic relationships of amphisbaenians, we performed a phylogenetic analysis on a combined data set of morphological characters and nuclear gene sequences (*rag-1*, *c-mos*) for extant and fossil squamates using parsimony and Bayesian methods (Fig. 3a). Analyses of combined data recover a monophyletic lacertid-amphisbaenian ('lacertibaenian'²⁴) clade, with *Cryptolacerta* clustering as sister taxon to Amphisbaenia in both the parsimony and Bayesian analyses (Fig. 3a). The sister relationship with Amphisbaenia is supported by 19 characters distributed across the entire skeleton (see Supplementary Information). Although homoplasy is common in many squamate osteological characters^{9,18}, the tongue-and-groove articulation of the frontals is unique to *Cryptolacerta* and Amphisbaenia, the transversely widened frontal downgrowths occur

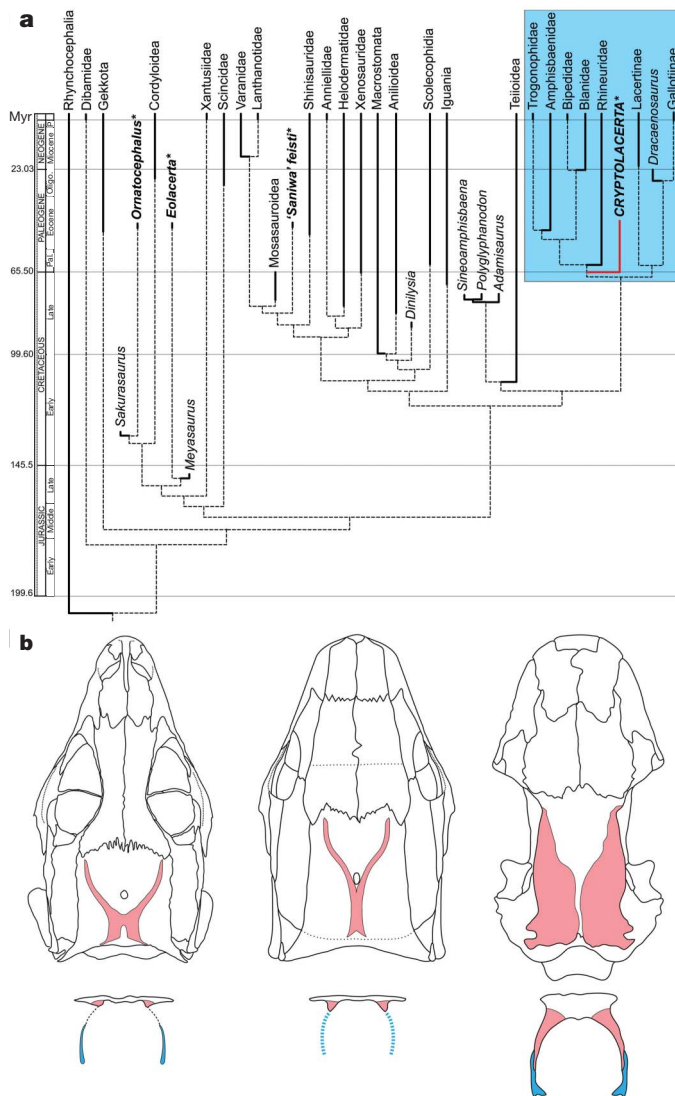


Figure 3 | Phylogeny of *Cryptolacerta* and the evolution of cranial akinesis in the origin of the amphisbaenian skull. **a**, Time-calibrated phylogeny of Squamata based on Bayesian analysis of morphological and molecular characters. Bold taxa with asterisks are from the Messel locality; the blue box denotes Lacertibaenia. **b**, Evolution of the amphisbaenian skull. Dorsal skull roofs and transverse sections through the dorsal braincase of (from left to right) *Podarcis pityusensis* (Lacertidae), *Cryptolacerta hassiaca* gen. et sp. nov. and *Rhineura floridana* (Amphisbaenia)³⁰. The ventroparietal crest of lacertids is ventrally connected with the prootic (blue) through a membranous sheet. In *Cryptolacerta* the crest is more prominent and must have had an extensive contact with the prootic. In amphisbaenians the crest is in full contact with the prootic and forms a secondary temporal region.

otherwise only in some scolecophidian snakes, the thickened frontal and maxilla can be found to a variable extent only in dibamids, snakes and varanids, the absence of a tympanic crest and the very low neural spines otherwise occur only within fossorial snakes, dibamids and some anguimorphs, and a sutural prefrontal–postorbitofrontal contact is shared only by some anguimorphs and *Sineoamphisbaena*. This last taxon, a Mesozoic squamate previously considered a stem-amphisbaenian¹⁰, falls within Polyglyphanodontidae as in other recent analyses¹¹ (Fig. 3a).

Cranial osteology of *Cryptolacerta* provides the first evidence of the origin of the derived amphisbaenian skull (Fig. 3b). In both *Cryptolacerta* and amphisbaenians the skull is reinforced by a strong vertical interdigitation of the frontals, thick, dense maxillae and frontals, and ventral downgrowth of the parietal. In lacertids the anterolateral portions of the ventroparietal crest closely approach the membranous alar processes of the prootic, and in amphisbaenians a membranous extension of the prootic is sutured to the ventrally extending parietal (Fig. 3b). Although the crest in *Cryptolacerta* is similar to lacertids, it is much more strongly developed and we infer extensive contact between the parietal and prootic. Additionally in basal amphisbaenians, the dorsal outline of the parietal table strongly reflects the shape of the ventroparietal crest of *Cryptolacerta*^{13–15}, suggesting that the lateral parts of a lacertid-like parietal became reduced during amphisbaenian evolution.

Body shape in squamates corresponds to locomotory habits^{19,20}, and the nearly complete skeleton of *Cryptolacerta* provides an opportunity to infer ecology near the origin of Amphisbaenia. To determine the habits of the taxon, we morphometrically analysed body shape in *Cryptolacerta* and extant squamates occupying habitats represented in the Messel depositional system (Fig. 4). Principal component analysis of cranial, axial and appendicular measurements¹⁹ produced morphospaces within which ecological habits were defined for extant taxa, and inferred for *Cryptolacerta*. For all coordinated principal component axes, *Cryptolacerta* falls outside the morphospace defined by fully fossorial squamates (Fig. 4). Although the reinforced skull and superficially small limbs are suggestive of fossorial habits, *Cryptolacerta* occupies a position within morphospace defined by taxa that are cryptic, leaf litter specialists and opportunistic burrowers (Fig. 4), based on relative body size, limb lengths and head size.

Ecomorphometry of *Cryptolacerta* and adaptations for a reinforced skull indicate that the early ecology of amphisbaenians and their relatives consisted of cryptic behavioural habits combined with head-first substrate locomotion, possibly as a defensive or predation strategy.

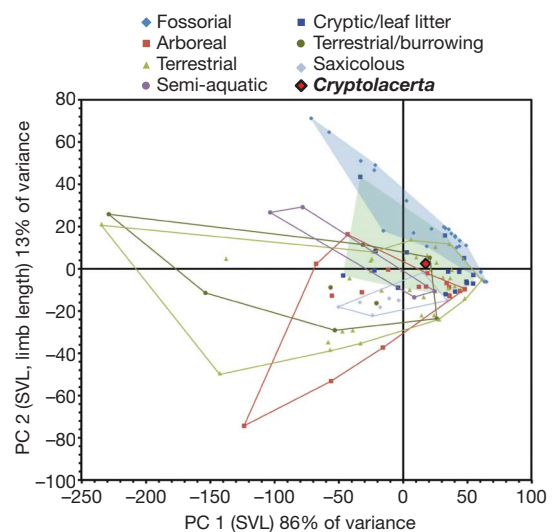


Figure 4 | Ecomorphology of *Cryptolacerta*. Principal component analysis of squamate morphology with ecological habits projected into shape space. Fossorial and cryptic morphospaces are shaded. *Cryptolacerta* occupies a position within the cryptic and terrestrial habit spaces.

Body elongation and limb reduction or loss are often considered prerequisites of fossoriality in squamates^{21–23}; however, *Cryptolacerta* demonstrates that modifications to cranial architecture preceded post-cranial specializations in amphisbaenians, showing that hypotheses of ecological character correlation may mask radically different histories of character evolution in ecological specialization^{18,23}.

Recent molecular divergence estimates²⁴ and the fossil record²⁵ indicate that lacertids and amphisbaenians diverged in the Late Cretaceous, at least 20 Myr before the occurrence of *Cryptolacerta*. The late occurrence of *Cryptolacerta* is consistent with hypotheses that intermediate body forms in the evolution of body elongation and limblessness can persist for tens of millions of years²⁰. It also suggests that the Palaeogene of Europe was a refugium for archaic Mesozoic squamate lineages, as indicated by co-occurring *Eolacerta*, '*Saniwa*', and *Ornatocephalus*^{9,26–28} (Fig. 3a), probably resulting from the island geography of Europe during the Late Cretaceous and early Cenozoic²⁹.

METHODS SUMMARY

The specimen was scanned by CT at the Helmholtz Centre Berlin for Materials and Energy using a micro-focus X-ray tube. Cranial reconstructions were performed using a wax model based on the CT data. Phylogenetic analyses were run using a partitioned data set of 364 morphological and 3,216 molecular (*rag-1*, *c-mos*) characters as well as 65 terminal taxa, using both parsimony and Bayesian methodology. Morphometric analysis used principal component analysis of a published data set of linear measurements of squamate body forms for taxa inhabiting environments represented in the Grube Messel depositional system¹⁹ to which *Cryptolacerta* was added based on measurements obtained from a digital calliper.

Received 22 January 2011; accepted 9 February 2011.

- Kearney, M. Systematics of the Amphisbaenia (Lepidosauria: Squamata) based on morphological evidence from recent and fossil forms. *Herpet. Monog* **17**, 1–74 (2003).
- Gans, C. & Montero, R. in *Biology of the Reptilia*, Vol. 21 (eds Gans C., Gaunt A. S. & Adler, K.) 621–738 (Society for the Study of Amphibians and Reptiles, 2008).
- Townsend, T. M., Larson, A., Louis, E. & Macey, J. R. Molecular phylogenetics of Squamata: the position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. *Syst. Biol.* **53**, 735–757 (2004).
- Vidal, N. & Hedges, S. B. The phylogeny of squamate reptiles (lizards, and amphisbaenians) inferred from nine protein-coding genes. *C. R. Biol.* **328**, 1000–1008 (2005).
- Estes, R., de Queiroz, K. & Gauthier, J. A. in *Phylogenetic Relationships of the Lizard Families* (eds Estes, R. & Pregill, G. K.) 119–281 (Stanford Univ. Press, 1988).
- Rieppel, O. A review of the origin of snakes. *Evol. Biol.* **22**, 37–130 (1988).
- Hallerman, J. The ethmoidal region of *Dibamus taylori* (Squamata: Dibamidae), with a phylogenetic hypothesis on dibamid relationships within Squamata. *Zool. J. Linn. Soc.* **122**, 385–426 (1998).
- Evans, S. & Barbadillo, L. J. An unusual lizard from the Early Cretaceous of Las Hoyas, Spain. *Zool. J. Linn. Soc.* **124**, 235–265 (1998).
- Conrad, J. L. Phylogeny and systematics of Squamata (Reptilia) based on morphology. *Bull. Am. Mus. Nat. Hist.* **310**, 1–182 (2008).
- Wu, X.-C. *et al.* Oldest known amphisbaenian from the Upper Cretaceous of Chinese Inner Mongolia. *Nature* **366**, 57–59 (1993).
- Kearney, M. The phylogenetic position of *Sineoamphisbaena hexatubularis* reexamined. *J. Vertebr. Paleontol.* **23**, 394–403 (2003).
- Franzen, J. L., Weber, J. & Wuttke, M. Senckenberg-Grabungen in der Grube Messel bei Darmstadt. 3. Ergebnisse 1979–1981. *Cour. Forsch.-Inst. Senckenberg* **54**, 12–15 (1982).
- Berman, D. S. *Spathorhynchus fossorium*, a middle Eocene amphisbaenian (Reptilia) from Wyoming. *Copeia* **1973**, 704–721 (1973).
- Berman, D. S. *Spathorhynchus natronicus*, a new species of rhineurid amphisbaenian (Reptilia) from the Early Oligocene of Wyoming. *J. Paleontol.* **51**, 986–991 (1977).
- Kearney, M., Maisano, J. A. & Rowe, T. Cranial anatomy of the extinct amphisbaenian *Rhineura hatcherii* (Squamata, Amphisbaenia) based on high-resolution X-ray computed tomography. *J. Morphol.* **264**, 1–33 (2005).
- Lee, M. S. Y. Hidden support from unpromising data sets strongly unites snakes with anguimorph 'lizards'. *J. Evol. Biol.* **22**, 1308–1316 (2009).
- Wiens, J. J. *et al.* Combining phylogenomics and fossils in higher level squamate reptile phylogeny: molecular data change the placement of fossil taxa. *Syst. Biol.* **59**, 675–688 (2010).
- Lee, M. S. Y. Convergent evolution and character correlation in burrowing reptiles: towards a resolution of squamate relationships. *Biol. J. Linn. Soc.* **65**, 369–453 (1998).
- Wiens, J. J., Brandley, M. C. & Reeder, T. W. Why does a trait evolve multiple times within a clade? Repeated evolution of snakelike body form in squamate reptiles. *Evolution* **60**, 123–141 (2006).
- Brandley, M. C., Huelsenbeck, J. P. & Wiens, J. J. Rates and patterns in the evolution of snake-like body form in squamate reptiles: evidence for repeated re-evolution of lost digits and long-term persistence of intermediate body forms. *Evolution* **62**, 2042–2064 (2008).
- Gans, C. Tetrapod limblessness: evolution and functional corollaries. *Am. Zool.* **15**, 455–467 (1975).
- Greer, A. E. Limb reduction in squamates: identification of the lineages and discussion of the trends. *J. Herpetol.* **25**, 166–173 (1991).
- Wiens, J. J. & Slingluff, J. L. How lizards turn into snakes: a phylogenetic analysis of body-form evolution in anguid lizards. *Evolution* **55**, 2303–2318 (2001).
- Hipsley, C. A., Himmelmann, L., Metzler, D. & Müller, J. Integration of Bayesian molecular clock methods and fossil-based soft bounds reveals early Cenozoic colonization of African lacertid lizards. *BMC Evol. Biol.* **9**, 151 (2009).
- Estes, R. *Handbuch der Paläoherpetologie* Pt 10 (Gustav Fischer, 1983).
- Müller, J. Osteology and relationships of *Eolacerta robusta*, a lizard from the middle Eocene of Germany (Reptilia, Squamata). *J. Vertebr. Paleontol.* **21**, 261–278 (2001).
- Rieppel, O., Conrad, J. L. & Maisano, J. A. New morphological data for *Eosaniwa koehni* Haubold 1977, and a revised phylogenetic analysis. *J. Paleontol.* **81**, 760–769 (2007).
- Weber, S. *Ornatocephalus metzleri* gen. et spec. nov. (Lacertilia, Scincoida) – taxonomy and paleobiology of a basal scincoid lizard from the Messel formation (middle Eocene: basal Lutetian, Geiseltalium), Germany. *Abh. Senckenb. Naturf. Gesellsch.* **561**, 1–159 (2004).
- Popov, S. V. *et al.* Lithological-Paleogeographic maps of the Paretethys. 10 maps Late Eocene to Pliocene. *Cour. Forsch.-Inst. Senckenberg* **250**, 1–46 (2004).
- The Deep Scaly Project. *Rhineura floridana*, Florida worm lizard. (http://digimorph.org/specimens/Rhineura_floridana/) (Digimorph, 2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Schaal for making the specimen available for study, J. J. Wiens for providing his morphometric data set of extant squamates, J. L. Conrad for providing his morphological data set of fossil and extant squamates, I. Manke for technical and administrative support, H. Mewis, H. Stöhr and A. Paulke for technical support, C. Bell and D. Evans for discussion, and P. Holroyd, M. Kroniger, M. Cunningham, M.-O. Rödel, K. Seymour, H.-D. Sues and F. Tillack for access to specimens. J.J.H. was funded by Natural Sciences and Engineering Research Council of Canada Discovery and Research Tools and Instruments grants, R.R.R. was funded by a Natural Sciences and Engineering Research Council of Canada Discovery grant and the Alexander von Humboldt-Stiftung.

Author Contributions J.M., C.A.H., J.J.H. and R.R.R. contributed to project planning and anatomical analysis. J.M., C.A.H., N.K. and A.H. contributed to micro-CT scanning. J.M. contributed to phylogenetic analysis. J.J.H. contributed to ecomorphological analysis. M.W. contributed to project initiation. All authors contributed to manuscript and figure preparation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.M. (johannes.mueller@mf-n-berlin.de).

Species–area relationships always overestimate extinction rates from habitat loss

Fangliang He^{1,2} & Stephen P. Hubbell^{3,4}

Extinction from habitat loss is the signature conservation problem of the twenty-first century¹. Despite its importance, estimating extinction rates is still highly uncertain because no proven direct methods or reliable data exist for verifying extinctions. The most widely used indirect method is to estimate extinction rates by reversing the species–area accumulation curve, extrapolating backwards to smaller areas to calculate expected species loss. Estimates of extinction rates based on this method are almost always much higher than those actually observed^{2–5}. This discrepancy gave rise to the concept of an ‘extinction debt’, referring to species ‘committed to extinction’ owing to habitat loss and reduced population size but not yet extinct during a non-equilibrium period^{6,7}. Here we show that the extinction debt as currently defined is largely a sampling artefact due to an unrecognized difference between the underlying sampling problems when constructing a species–area relationship (SAR) and when extrapolating species extinction from habitat loss. The key mathematical result is that the area required to remove the last individual of a species (extinction) is larger, almost always much larger, than the sample area needed to encounter the first individual of a species, irrespective of species distribution and spatial scale. We illustrate these results with data from a global network of large, mapped forest plots and ranges of passerine bird species in the continental USA; and we show that overestimation can be greater than 160%. Although we conclude that extinctions caused by habitat loss require greater loss of habitat than previously thought, our results must not lead to complacency about extinction due to habitat loss, which is a real and growing threat.

The Millennium Ecosystem Assessment¹ predicts that near-term extinction rates could be as high as 1,000 to 10,000 times background rates (see also ref. 7). Most predictions of species extinction rates, including those in the Millennium Ecosystem Assessment, are inferred from applying the SAR to rates of habitat loss^{8–14}. The wide discrepancy between the rates of species extinction predicted by this method and the extinction rates actually recorded, has fuelled a continuing debate about how to explain the discrepancy^{2,4,15–20}. The main issue is that, almost always, more species are left after a given loss of habitat than the number of species predicted to remain, based on the SAR. The most frequent interpretation is that the excess species are ‘committed to extinction’. The term ‘extinction debt’ was coined to refer to species’ populations that were no longer viable but were facing certain extinction due to habitat destruction that had already occurred^{3,6,17}. The consensus on the most likely reason for the extinction debt is that there is a time lag for populations to go extinct after severe losses in population size^{6,21}.

Here we show that extinction rates estimated from the SAR are all overestimates. We define extinction rate as the fractional loss of species over a defined period accompanied by a given loss of habitat. These overestimates are due to the false assumption that the sampling problem for extinction is simply the reverse of the sampling problem for the SAR. The area that must be added to find the first individual of a species is in

general much smaller than the area that must be removed to eliminate the last individual of a species (Fig. 1). Therefore, on average, it takes a much greater loss of area to cause the extinction of a species than it takes to add the species on first encounter, except in the degenerate case of a species having a single individual. We show mathematically that this is a necessary result of fundamental sampling differences between the SAR and the endemics–area relationship (EAR). Only in a very special and biologically unrealistic case, when all species are randomly and independently distributed in space, is it possible to derive the EAR from the SAR. Although this special case almost never occurs in nature, we examine this simple case first to clarify the nature of the problem. Then we relax these assumptions and consider the general case of aggregated species distributions.

The problem has gone unnoticed for so long because the traditional method for estimating extinction uses the power-law SAR, $S = cA^z$, which has no sampling theory relating it to species distributions (Supplementary Information A). To develop a sampling theory, we must consider the spatial distribution of species explicitly (Supplementary Information B and C). We derive the SAR and EAR from nearest-neighbour distances under two situations, random dispersion and clumped dispersion. We construct a SAR from the probability of encountering the first nearest neighbour of a species (a new species is added every time the sampling frame a encounters the first individual of the given species). In contrast, we construct the EAR from the probability of encountering the last neighbour of a species (a species is added only after all individuals are contained within frame a). We arrive at the species–area curve for randomly and independently distributed species as (Supplementary Information B):

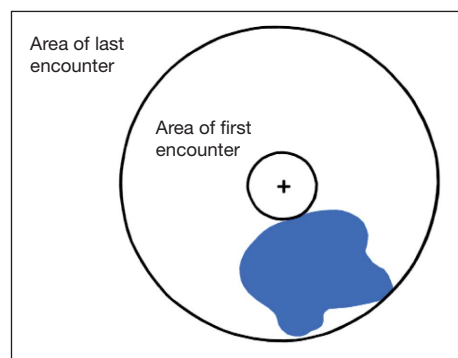


Figure 1 | Sampling differences for SAR and EAR. Range distribution of a species (blue area), and an arbitrary starting sample point, indicated by +. Regardless of the starting location, a sampling frame of arbitrary shape (here circular) with an area of a size sufficient to contact the species for the first time is always less than the sample area needed to encompass the entire range of the species. The SAR (species accumulation) is constructed from sample areas of first contact, and the EAR (species extinction) is constructed from areas of last contact.

¹State Key Laboratory of Biocontrol and School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China. ²Department of Renewable Resources, University of Alberta, Edmonton, Alberta, T6G 2H1, Canada. ³Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California 90095, USA. ⁴Center for Tropical Forest Science, Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002-0948, Republic of Panama.

$$S_a^1 = S - \sum_{i=1}^S \left(1 - \frac{a}{A}\right)^{N_i} \quad (1)$$

and the endemics–area curve as:

$$S_a^N = \sum_{i=1}^S \left(\frac{a}{A}\right)^{N_i} \quad (2)$$

where N_i is the total abundance of species i and S is the total number of species in the region A . Equations (1) and (2), derived from nearest-neighbour distances, are identical to the classical random placement models^{22–25}.

Let the total area be A and let a sub-area a be lost. For randomly and independently distributed species, we can calculate the expected number of species lost with a loss of area a from the SAR (equation (1)) as $S_{\text{loss}} = S - S_{A-a}$. This is identical to the EAR calculated directly from equation (2): $S_{\text{loss}} = S - S_{A-a}^1 = \sum_{i=1}^S \left(\frac{a}{A}\right)^{N_i} = S_a^N$. This proves that, for the special case of species distributed randomly in space, extinction rates estimated from the backward random placement SAR and from the forward random placement EAR are the same, and the SAR and EAR are mirror images (Fig. 2 and Supplementary Fig. 1). This case is true because, under random placement, the total area A is equal to the sum of the areas of encountering the first individual and the last individual of a species. From the probability models of the nearest-neighbour distance, the expected area needed to sample the first individual is $a^1 = A/(N+1)$, and the expected area for the last individual is $a^N = NA/(N+1)$ (Supplementary Information B). Thus $a^1 + a^N = A$. Note that $a^N > a^1$ is always true except when $N = 1$.

This mirror-image relationship only holds for randomly distributed species, however. Almost all species in nature are clumped, not randomly distributed²⁶. For aggregated species, one can show that $a^1 + a^N < A$ with $a^N \geq a^1$ remaining true (Supplementary Information C and Supplementary Fig. 2). This leads to $S - S_{A-a}^1 \neq S_a^N$. The more spatially aggregated species distributions are, the stronger the inequality $a^N \geq a^1$ becomes. These results are completely general and explain the discrepancy between the backward SAR and forward EAR methods as well as why the backward SAR method systematically overestimates extinction rates.

These results apply to sample areas on any spatial scale. We can assess the magnitude of overestimation by the backward SAR method precisely in cases where we know the species composition and spatial location of each individual of each species or spatial range of each species. To illustrate this, we use spatially explicit data from eight large stem-mapped plots from a global forest dynamics network. We also perform the analysis on biogeographical spatial scales for passerine species in the continental USA (see Methods). The results show that the classic power-law SAR model, $S = cA^z$, and its corresponding EAR model (Supplementary Information A),

$$\lambda = S_{\text{loss}}/S_A = 1 - (1 - a/A)^z \quad (3)$$

are not mirror-image curves. In equation (3), S_{loss} is the number of species lost (endemic) to destroyed sub-area a . Because of the difference in sampling procedure of encountering species and losing species, the slopes z of the power-law model $S = cA^z$ and EAR (3) are not the same. The fit of the power-law SAR and EAR to species–area and endemics–area data respectively lead to two very different slopes

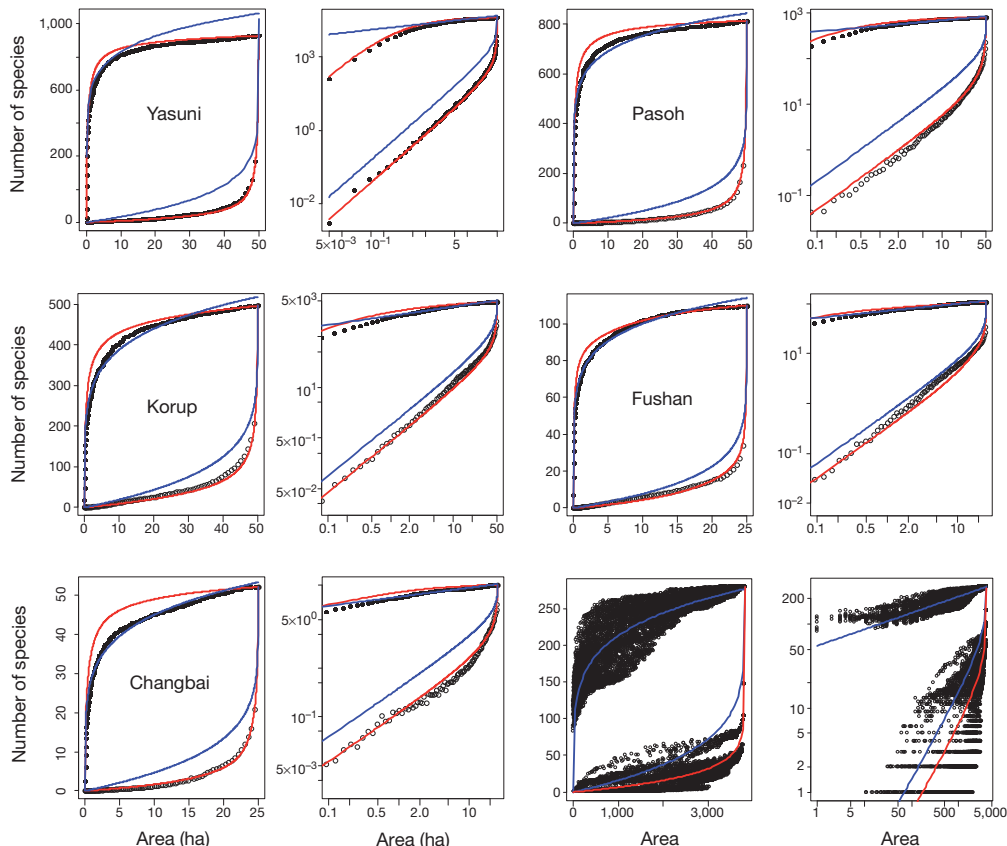


Figure 2 | Species– and endemics–area curves for six of the nine data sets in Table 1. The second and fourth columns are the plots on a log–log scale. The upper and lower blue curves are the fits of the power-law SAR and EAR (equation (3)), respectively. The upper and lower red curves are the predictions of the random placement SAR (equation (1)) and EAR (equation (2)), respectively. Unlike for the other data sets, the red curve for US passerine data

(cell size 0.48° latitude \times 0.48° longitude) is the fit of equation (3) because the abundances of the passerine species are not known (so equation (2) cannot be used). The cloud of points represent 100 repeated random samples of the SAR and EAR. The SAR and EAR curves for the Barro Colorado Island plot are shown in Supplementary Fig. 1.

Table 1 | Eight stem-mapped forest plots across the world and distributions of passerine birds in the continental USA

Plot	Forest type	Size (ha)	Number of trees	Number of species	z_{SAR}	z_{EAR}	Bias _{0.52} (%)	Bias ₂₅ (%)
Barro Colorado Island, Panama	Lowland tropical forest	50	325,549	316	0.133 (0.00202)	0.0803 (0.000611)	65.61	64.38
Yasuni, Ecuador	Lowland tropical forest	50	307,279	1,128	0.126 (0.00473)	0.0623 (0.00189)	102.21	100.41
Pasoh, Malaysia	Lowland tropical forest	50	323,262	814	0.124 (0.00374)	0.0536 (0.00158)	131.30	129.02
Korup, Cameroon	Lowland tropical forest	50	328,973	496	0.179 (0.00369)	0.113 (0.00116)	58.38	56.92
Dinghu, China	Subtropical evergreen broad-leaved forest	20	71,617	210	0.274 (0.00180)	0.193 (0.000880)	41.94	40.34
Fushan, Taiwan	Subtropical evergreen broad-leaved forest	25	114,508	110	0.142 (0.00199)	0.0922 (0.000838)	53.99	52.92
Tiantong, China	Subtropical evergreen broad-leaved forest	20	94,603	152	0.200 (0.00214)	0.0994 (0.00175)	101.15	98.34
Changbai, China	Temperate forest	25	38,902	52	0.184 (0.00296)	0.0905 (0.00233)	103.27	100.62
USA	Passerine birds (0.24° × 0.24°)	14,904	–	279	0.187 (0.00101)	0.0766 (0.000516)	144.06	140.31
USA	Passerine birds (0.48° × 0.48°)	3,830	–	279	0.195 (0.00106)	0.0791 (0.000421)	147.39	143.39

The 'bias' is the overestimation calculated by comparing the extinction rates estimated from the z_{SAR} values with those from the endemic z_{EAR} values: $(\lambda_{\text{SAR}} - \lambda_{\text{EAR}})/\lambda_{\text{EAR}}$. We calculated percentage bias by assuming 0.52% and 25% habitat loss^{1–29}, respectively. Equation (3) gives λ_{SAR} . To analyse passerine distributions, we divided the lower 48 states of the USA into a grid of 14,904 cells with cell size of 0.24° latitude × 0.24° longitude and into 3,830 cells with cell size of 0.48° latitude × 0.48° longitude.

(the SAR z_{SAR} versus the EAR z_{EAR}) (Table 1). In some cases, z_{SAR} can be more than double z_{EAR} . This result is independent of the spatial scale of the data, as is evident for the passerine case shown in Table 1.

This analysis demonstrates that the most widely used method of estimating species extinction rates due to habitat loss, the backward SAR calculation, is not correct. For non-randomly distributed species, the SAR and EAR are not mirror images, so that one cannot be used to infer the other. This result holds regardless of how well the power-law SAR fits species–area data (Supplementary Information D). Even for randomly distributed species, the backward power-law SAR model is still not appropriate for estimating extinction rates because in this case equation (1) is the only correct SAR, not the power-law SAR (Supplementary Fig. 2), and equation (2) is the only correct EAR model. These results show that the concept of an 'extinction debt' (that is, the extinctions lost to biotic relaxation due to habitat destruction) based on the backward SAR model is not conceptually sound. Note that these results say nothing about whether an extinction debt exists, only that such a debt as might exist is not appropriately measured by the backward SAR method. To model the process of biotic relaxation will require a dynamic theoretical framework different from the current static SAR model. Currently, no such theory is available. The EAR curve is consistent with the concept of 'imminent extinction', which states that predictions of near-term extinctions due to habitat loss should focus on species endemic to the area of destroyed habitat^{4,7,27,28}.

Previous estimates of extremely high extinction rates, – for example, one species per hour to one species a day⁸, 33–50% of all species between the 1970s and 2000 (ref. 9), from half to several million species by 2000 (refs 10, 12) or 50% of species by 2000 (ref. 11) – have not been observed. There is also reason to question the recent estimates of extinction rates made by the Millennium Ecosystem Assessment¹ and those by Thomas *et al.*¹⁹. In the latter case, the loss of habitat and the shift of species' ranges are driven by climate change. However, the use of the flawed backward SAR in Thomas *et al.* raises a legitimate question about the validity of their conclusion that 18–35% of species will be committed to extinction by 2050. We suggest that their estimated rates of extinction should be regarded as a high-end possibility rather than as supported by hard scientific evidence.

By how much have we overestimated extinction rates? Precise answers to this question require information about the EAR curve, which is generally not known. However, we can make a first approximation from the results shown in Table 1, for which we know the EAR curves in stem-mapped samples of forests and range distributions of passerines. We calculated the z_{EAR} and z_{SAR} averaged over the data in Table 1, leading to $z_{\text{EAR}} = 0.0940$ and $z_{\text{SAR}} = 0.174$. We then used two estimates of forest habitat loss, the annual deforestation rate of $(a/A)100\% = 0.52\%$ for humid tropical forests²⁹ and the estimated 25% conversion of original forest habitat into agricultural land¹. The

SAR backward method (equation (3)) overestimates extinction rates by 85.07% and 83.00% in these respective cases, compared with the rates estimated by the forward EAR method. Conservation biologists often use a z value of 0.25 in cases where z values are not available^{18,19}. Using this value inflates extinction rate estimates much more, being 165.85% and 160.10% for the two deforestation rates, respectively (see also Supplementary Fig. 1).

Are better methods available for estimating extinction rates? Our results show that the random placement EAR curve describes the empirical EAR curves for the forest plots very well. This result is remarkable and provides a simple method for estimating extinction (Supplementary Information E). Note that the theoretical random placement EAR for each plot is not data-fitting but a genuine prediction from equation (2).

These results might receive a mixed reaction from the conservation community. On the one hand, the good news is that all extinction rate estimates based on the backward SAR method are overestimates. Because it is derived from sample areas of first contact with each species, the backward SAR method makes the previously unrecognized assumption that any loss whatsoever of population due to habitat loss commits a species to extinction, which clearly is not true. On the other hand, there is likely to be concern that these results could jeopardize conservation efforts and be falsely construed in some quarters to imply that habitat loss is not a problem. Nothing could be further from the truth. There is no doubt whatsoever that the Millennium Ecosystem Assessment¹ has correctly identified habitat loss as the primary threat to conserving the Earth's biodiversity, and the sixth mass extinction might already be upon us or imminent³⁰. Our results do indicate, however, that the backward SAR is not the correct way to estimate the magnitude of the current extinction event. To help mitigate contemporary extinctions and strengthen the science behind conservation planning, we need far better geographical data on endemism and species' distributions to improve forecasts of extinction rates⁷. Improving geographical databases on the distribution of biodiversity on Earth should be an urgent international priority.

METHODS SUMMARY

We analysed data from eight 20–50 ha (1 hectare (ha) = 10⁴ m²), stem-mapped plots of the Center for Tropical Forest Science global plot network to construct SAR and EAR curves (<http://www.ctfs.si.edu/>). These data sets are suitable because (1) our analysis is independent of spatial scale, (2) they are among the few data sets in which individuals are mapped on a landscape scale and (3) the EAR curve, which must be known, cannot be calculated from SAR curves (see text).

We obtained the SAR and EAR curves as follows: (1) grid the plot into cells of some minimum size (for example 5 m × 5 m); (2) count the number of species and the number of endemic species (species completely confined to the sample area) in each cell; (3) average the number of species per cell and the number of endemic species across all cells of a given size; and (4) construct species–area and endemics–area

curves by repeating steps 1–3, increasing cell size from $5\text{ m} \times 5\text{ m}$, $10\text{ m} \times 5\text{ m}$, $10\text{ m} \times 10\text{ m}$, etc. up to the entire plot.

We estimated z_{SAR} by nonlinear fit of the power-law SAR model to the observed species–area data. We limited fitting to areas of at least 0.2 ha because the power-law model is not considered applicable at small spatial scales (including them inflates z values and worsens overestimation). We estimated z_{EAR} by directly fitting equation (3) to the observed endemics–area data (see Table 1).

We analysed SAR and EAR curves for 279 passerine species in the lower 48 states of the USA using individual species' range maps from Natureserve (<http://www.natureserve.org/getData/birdMaps.jsp>). We divided the USA into grids at two respective cell sizes, 0.24° latitude \times 0.24° longitude (14,904 cells) and 0.48° latitude \times 0.48° longitude (3,830 cells), to confirm that our analysis is robust to scale change, as predicted by the analytical results. We computed SAR and EAR curves using presence–absence data following the above procedure.

Received 21 December 2010; accepted 8 March 2011.

- Millennium Ecosystem Assessment. *Ecosystems and Human Well-Being: Biodiversity Synthesis* (World Resources Institute, 2005).
- Simberloff, D. in *Tropical Deforestation and Species Extinction* (eds Whitmore, T. C. & Sayer, J. A.) 75–89 (Chapman & Hall, 1992).
- May, R. M., Lawton, J. H. & Stork, N. E. in *Extinction Rates* (eds Lawton, J. H. & May, R. M.) 1–24 (Oxford Univ. Press, 1995).
- Pimm, S. L. & Askins, R. A. Forest losses predict bird extinctions in eastern North America. *Proc. Natl Acad. Sci. USA* **92**, 9343–9347 (1995).
- Rosenzweig, M. L. *Species Diversity in Space and Time* (Cambridge Univ. Press, 1995).
- Tilman, D., May, R. M., Lehman, C. L. & Nowak, M. A. Habitat destruction and the extinction debt. *Nature* **371**, 365–366 (1994).
- Pimm, S. L., Russell, G. J., Gittleman, J. L. & Brooks, T. M. The future of biodiversity. *Science* **269**, 347–350 (1995).
- Myers, N. *The Sinking Ark: A New Look at the Problem of Disappearing Species* (Pergamon, 1979).
- Lovejoy, T. E. in *The Global 2000 Report to the President: Entering the Twenty-First Century* (study director Barney, G. O.) 328–331 (Council on Environmental Quality, U.S. Government Printing Office, 1980).
- National Research Council. *Research Priorities in Tropical Biology* (National Academy of Sciences, 1980).
- Ehrlich, P. R. & Ehrlich, A. H. *Extinction: The Causes and Consequences of the Disappearance of Species* (Random House, 1981).
- Myers, N. in *Tropical Rain Forest Ecosystems, Structure and Function* (ed. Golley, F. B.) 325–334 (Elsevier, 1983).
- Lugo, A. E. in *Biodiversity* (ed. Wilson, E. O.) 58–70 (National Academy Press, 1988).
- WCMC. *Global Biodiversity: Status of the Earth's Living Resources* (Chapman & Hall, 1992).
- Heywood, V. H. & Stuart, S. N. in *Tropical Deforestation and Species Extinction* (ed. Whitmore, T. C. & Sayer, J. A.) 91–117 (Chapman & Hall, 1992).
- Dial, R. & Budiansky, S. Extinction or miscalculation? *Nature* **370**, 104–105 (1994).
- Heywood, V. H., Mace, G. M., May, R. M. & Stuart, S. N. Uncertainties in extinction rates. *Nature* **368**, 105 (1994).
- Brooks, T. & Balmford, A. Atlantic forest extinctions. *Nature* **380**, 115 (1996).
- Thomas, C. D. *et al.* Extinction risk from climate change. *Nature* **427**, 145–148 (2004).
- Harte, J., Ostling, A., Green, J. & Kinzig, A. Biodiversity conservation: Climate change and extinction risk. *Nature* **430**, doi:10.1038/nature02718 (2004).
- Brooks, T. M., Pimm, S. L. & Oyugi, J. O. Time lag between deforestation and bird extinction in tropical forest fragments. *Conserv. Biol.* **13**, 1140–1150 (1999).
- Arrhenius, O. Species and area. *J. Ecol.* **9**, 95–99 (1921).
- Coleman, B. D. Random placement and species–area relations. *Math. Biosci.* **54**, 191–215 (1981).
- He, F. & Legendre, P. Species diversity patterns derived from species–area models. *Ecology* **83**, 1185–1198 (2002).
- Green, J. & Ostling, A. Endemics–area relationships: the influence of species dominance and spatial aggregation. *Ecology* **84**, 3090–3097 (2003).
- Condit, R. *et al.* Spatial patterns in the distribution of tropical tree species. *Science* **288**, 1414–1418 (2000).
- Kinzig, A. P. & Harte, J. Implications of endemics–area relationships for estimates of species extinctions. *Ecology* **81**, 3305–3311 (2000).
- Pimm, S. L. & Raven, P. Extinction by numbers. *Nature* **403**, 843–845 (2000).
- Achard, F. *et al.* Determination of deforestation rates of the world's humid tropical forests. *Science* **297**, 999–1002 (2002).
- Barnosky, A. D. *et al.* Has the Earth's sixth mass extinction already arrived? *Nature* **471**, 51–57 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge the work of the principal investigators and their field assistants for collecting the field data on the large plots of the Center for Tropical Forest Science network. We are grateful to P. Gowaty, K. Gaston and M. Rosenzweig for their comments on the paper. We thank R. Condit, R. Perez and S. Aguilar for Barro Colorado Island, R. Valencia for Yasuni, N. Supardi Noor and R. Kassim for Pasoh, D. Thomas, D. Kenfack and G. Chuyong for Korup, W.-H. Ye for Dinghushan, I.-F. Sun and C.-F. Hsieh for Fushan, X.-H. Wang for Tiantong, Z.-Q. Hao for Changbaishan and X.-L. Wang for processing the passerine data. This work was supported by Sun Yat-sen University, the Natural Sciences and Engineering Research Council (Canada), NASA (National Aeronautics and Space Administration) and the US National Science Foundation.

Author Contributions Both authors designed the study, analysed and interpreted data, and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to F.H. (fhe@mail.syu.edu.cn; fhe@ualberta.ca) or S.P.H. (shubbell@eeb.ucla.edu).

Neurotrophin cleaves EphB2 in the amygdala to control anxiety

Benjamin K. Attwood¹, Julie-Myrtille Bourgognon¹, Satyam Patel¹, Mariusz Mucha¹, Emanuele Schiavon¹, Anna E. Skrzypiec¹, Kenneth W. Young², Sadao Shiosaka³, Michal Korostynski⁴, Marcin Piechota⁴, Ryszard Przewlocki⁴ & Robert Pawlak¹

A minority of individuals experiencing traumatic events develop anxiety disorders. The reason for the lack of correspondence between the prevalence of exposure to psychological trauma and the development of anxiety is unknown. Extracellular proteolysis contributes to fear-associated responses by facilitating neuronal plasticity at the neuron–matrix interface^{1–4}. Here we show in mice that the serine protease neurotrophin is critical for stress-related plasticity in the amygdala by regulating the dynamics of the EphB2–NMDA-receptor interaction, the expression of *Fkbp5* and anxiety-like behaviour. Stress results in neurotrophin-dependent cleavage of EphB2 in the amygdala causing dissociation of EphB2 from the NR1 subunit of the NMDA receptor and promoting membrane turnover of EphB2 receptors. Dynamic EphB2–NR1 interaction enhances NMDA receptor current, induces *Fkbp5* gene expression and enhances behavioural signatures of anxiety. On stress, neurotrophin-deficient mice do not show EphB2 cleavage and its dissociation from NR1 resulting in a static EphB2–NR1 interaction, attenuated induction of the *Fkbp5* gene and low anxiety. The behavioural response to stress can be restored by intra-amygdala injection of neurotrophin into neurotrophin-deficient mice and disrupted by the injection of either anti-EphB2 antibodies or silencing the *Fkbp5* gene in the amygdala of wild-type mice. Our findings establish a novel neuronal pathway linking stress-induced proteolysis of EphB2 in the amygdala to anxiety.

Fear helps organisms recognize, memorize and predict danger, thereby promoting their survival. However, severe stress can trigger maladaptive forms of neuronal remodelling leading to generalization of fear and high anxiety⁵.

Traumatic events are memorized as a result of the capacity of synaptic connections and the surrounding matrix to undergo experience-dependent functional or morphological changes^{1,6}. Extracellular proteases are strategically poised to remodel the neuron–extracellular-matrix interface and facilitate fear and anxiety^{2–4}. Eph-receptor tyrosine kinases constitute an important group of molecules subject to modulation by extracellular proteases⁷. Although Ephs promote neuronal plasticity^{8,9}, their involvement in behavioural responses to environmental stimuli is not clear.

Neurotrophin is a serine protease uniquely positioned to facilitate stress-induced plasticity due to its high expression in the amygdala and hippocampus¹⁰. To investigate if neurotrophin and Ephs co-localize we performed immunohistochemistry. Consistent with previous reports^{10,11} we found robust expression of both neurotrophin and EphB2 in the amygdala (Fig. 1 and Supplementary Fig. 1) and the hippocampus (not shown). Double immunohistochemistry revealed high levels of neurotrophin co-localizing with EphB2-rich clusters on amygdala neurons (Fig. 1a).

To assess whether Ephs are modulated by neurotrophin we treated SH-SY5Y cells with neurotrophin and measured the levels of Eph receptors by western blotting. We found that neurotrophin (but not other proteases;

Supplementary Fig. 2) cleaved EphB2 (decrease by 41%, $P < 0.001$), whereas the levels of other Ephs or their ligand, ephrinB2, remained unchanged (Fig. 2a, b and Supplementary Fig. 3a). When we expressed either GFP-tagged EphB2, GFP-tagged EphA4 or unlinked GFP in SH-SY5Y cells (Supplementary Figs 4, 5) and treated them with neurotrophin we saw a similar decrease in the EphB2-associated signal (Supplementary Fig. 5; $P < 0.05$).

When we used the above protocol to examine the composition of the SH-SY5Y or HEK293 cell culture medium after the application of neurotrophin, we found a new ~70 kDa extracellular fragment of EphB2 released into the media (Fig. 2c), the size of which was consistent with neurotrophin cleaving EphB2 close to the cell membrane (Supplementary

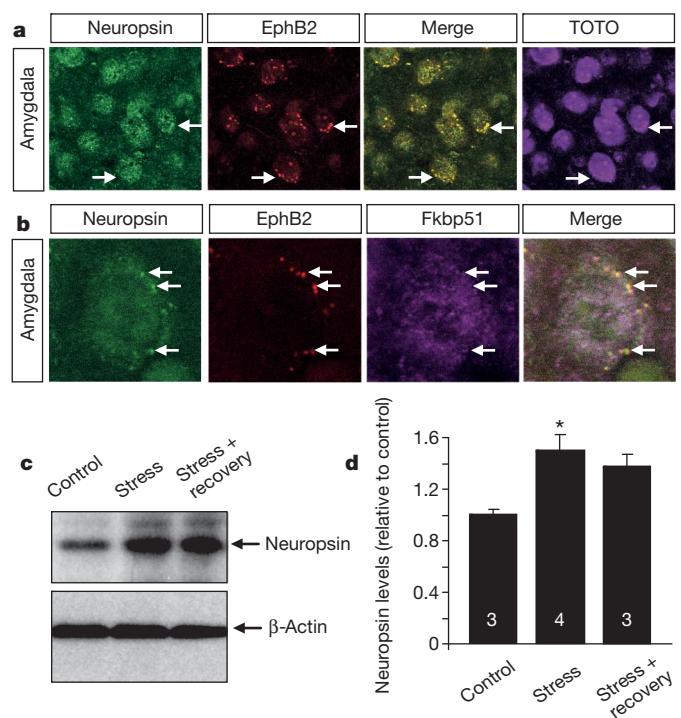


Figure 1 | Neurotrophin and EphB2 co-localize in neurons of the basolateral complex of the amygdala. **a**, Double immunohistochemistry showed neurotrophin (green) and EphB2 (red) co-localize in lateral amygdala neurons (arrows show EphB2-rich clusters at neurotrophin detection sites). Cells were highlighted with TOTO-3 stain. **b**, Triple immunohistochemistry confirmed the presence of neurotrophin/EphB2-rich clusters on the neuronal surface and a low degree of co-localization with cytoplasmic Fkbp51. **c**, **d**, Western blotting revealed amygdala neurotrophin upregulation after 6 h restraint stress ($F_{(2, 7)} = 8.81$; $P < 0.05$). Digits inside columns indicate n . * $P < 0.05$. Results are shown as mean \pm s.e.m.

¹Department of Cell Physiology and Pharmacology, University of Leicester, University Road, Leicester LE1 9HN, UK. ²Medical Research Council Toxicology Unit, Lancaster Road, Leicester LE1 9HN, UK. ³Division of Structural Cell Biology, Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan. ⁴Department of Molecular Neuropharmacology, Institute of Pharmacology, 12 Smetna Street 31-343 Krakow, Poland.

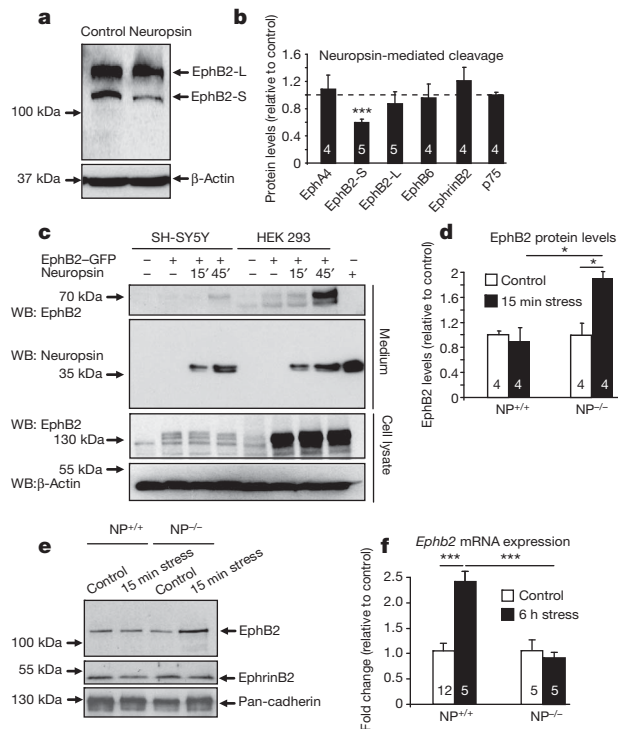


Figure 2 | Neuropilin cleaves EphB2 and regulates its expression both *in vitro* and in the amygdala after stress. **a**, **b**, EphB2-S (short splice variant) band density in SH-SY5Y cells decreased on 15 min of neuropilin treatment ($F_{(3,18)} = 11.24$; $P < 0.001$). Neuropilin did not cleave other molecules of the same class (**b** and Supplementary Fig. 3a). **c**, Exposure of EphB2-GFP-transfected SH-SY5Y or HEK293 cells to neuropilin (15 or 45 min) resulted in the appearance of a ~70 kDa amino-terminal EphB2 fragment in the medium (Supplementary Fig. 6). WB, western blot. **d**, **e**, A twofold increase in membrane-associated EphB2 in neuropilin^{-/-} ($F_{(3,12)} = 6.4$; $P < 0.05$ versus non-stressed) but not wild-type mice was observed after stress ($P < 0.05$ versus stressed neuropilin^{-/-} mice). **f**, RT-qPCR revealed a twofold upregulation of *Ephb2* gene expression after 6 h stress ($F_{(3,23)} = 13.48$; $P < 0.001$), not observed in neuropilin-deficient animals ($P < 0.001$ versus stressed wild-type mice). EphB2-S and EphB2-L describe short and long splice variants, respectively. NP, neuropilin. Digits inside columns indicate *n*. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Results are shown as mean \pm s.e.m.

Fig. 6a, b). Next, we subjected wild-type and neuropilin^{-/-} mice to restraint stress to activate the basolateral complex of the amygdala¹². Neuropilin levels increased by 50% after stress and gradually normalized during recovery in this brain region (Fig. 1c, d; $P < 0.05$). Western blotting revealed a twofold increase in membrane-associated amygdala EphB2 levels after 15 min of restraint stress in neuropilin^{-/-} mice (Fig. 2d, e and Supplementary Fig. 7; $P < 0.05$) indicative of new EphB2 receptors being incorporated into the membrane. This increase was not observed in wild-type mice, consistent with neuropilin-mediated EphB2 cleavage during stress. Cleavage of EphB2 in the amygdala of wild-type mice was followed by a twofold increase in the expression of the *Ephb2* gene (Fig. 2f; $P < 0.001$). The cleavage was substrate-specific because stress did not alter the levels of either ephrinB2 (Supplementary Fig. 9a) or a presynaptic neuropilin substrate, NCAM-L1 (also known as L1cam)¹³ (Supplementary Fig. 3b, c).

To determine the structural basis of neuropilin-specific EphB2 cleavage we analysed the fibronectin type III domain of EphB2 (Supplementary Fig. 8) looking for similarities with the previously published neuropilin cleavage sequences¹⁴. We found a critical amino acid pair, Gly-Arg, at position 517 of EphB2 but not EphB1, EphB6 or EphA4. Consistent with our experimental findings cleavage of EphB2 at this site would result in the release of a ~70 kDa extracellular fragment.

The EphB2 receptors cluster and associate with NMDA receptors at excitatory synapses^{15–17}. Indeed, co-immunoprecipitation revealed NR1 bound to EphB2 in the amygdala (Fig. 3a). Restraint stress reduced the amount of EphB2 associated with NR1 at 15 min by 42% (Fig. 3a, b; $P < 0.05$) whereas NR1 levels were not altered (Supplementary Fig. 9b–e). Stress-induced decrease in the EphB2–NR1 association was not observed in neuropilin^{-/-} mice but restored by intra-amygdala administration of neuropilin into these animals (Fig. 3a, b), consistent with neuropilin cleaving the extracellular portion of EphB2 during stress and triggering its dissociation from NR1. These results, together with stress-induced EphB2 membrane trafficking (Fig. 2d, e), indicate that neuropilin increases the dynamics of the EphB2–NR1 interaction after stress.

Regulating the EphB2–NMDA-receptor interaction results in modulation of the expression of NMDA-receptor-dependent genes facilitating synaptic plasticity¹⁷. To examine if neuropilin-mediated regulation of the EphB2–NR1 assembly affects gene expression in the amygdala we used microarrays in neuropilin^{-/-} and wild-type mice. We found 19 differentially expressed transcripts with a marked upregulation of the *Fkbp5* gene (Fig. 3c, d and Supplementary Figs 10, 11; $P < 0.0005$). This gene encodes the Fkbp51 protein, which regulates glucocorticoid receptor sensitivity. *Fkbp5* has been implicated in the development of anxiety, depression and post-traumatic stress disorder (PTSD)^{18–20}. Quantitative polymerase chain reaction with reverse transcription (RT-qPCR) confirmed an increase in *Fkbp5* gene expression in the amygdalae of stress-naïve neuropilin^{-/-} animals (Fig. 3e; $P < 0.05$).

The extent of upregulation of *Fkbp5* messenger RNA shortly after trauma correlates with the development of PTSD²¹. If neuropilin regulates *Fkbp5* gene expression then the magnitude of its stress-related regulation should be altered in neuropilin-deficient mice. When we analysed stress-induced *Fkbp5* gene expression we found a 21-fold upregulation in wild-type amygdalae (Fig. 3e; $P < 0.001$) but an attenuated upregulation in neuropilin^{-/-} mice. The increase in *Fkbp5* gene expression was accompanied by a twofold upregulation of Fkbp51 protein levels in wild-type mice but not neuropilin^{-/-} mice (Fig. 3f, g; $P < 0.05$). These results indicate that neuropilin is a key regulator of the *Fkbp5* gene and protein expression.

Neuropilin is an extracellular protease and thus unlikely to alter the expression of the *Fkbp5* gene directly. Although the *Fkbp5* gene can be regulated by glucocorticoids (Supplementary Fig. 12), the above differences in *Fkbp5* expression after stress cannot be attributed to corticosterone levels (Supplementary Fig. 13). Interference with EphB2 signalling has recently been linked to the regulation of the *Fkbp5* gene²². Indeed, when we mimicked stress *in vitro* by adding corticosterone into neuronal amygdala cultures, neuropilin-mediated upregulation of *Fkbp5* was hindered by anti-EphB2 antibody (Fig. 3h; $P < 0.001$) and imitated by NMDA receptor stimulation (Fig. 3i; $P < 0.05$).

To address the effect of neuropilin on NMDA receptors directly we measured the evoked NMDA/AMPA current ratio in principal neurons of the basal amygdala in wild-type and neuropilin^{-/-} mice. We found that, unlike in the hippocampus¹³, the NMDA current was markedly reduced by the deletion of the neuropilin gene, resulting in a ~50% drop in the NMDA/AMPA ratio (Fig. 4a–c; $P < 0.01$).

We next asked if the neuropilin pathway affects neuronal plasticity in the amygdala. We induced early (E-LTP) or sustained (L-LTP) long-term potentiation in the amygdala lateral-basal pathway of wild-type and neuropilin^{-/-} mice. Whereas basal synaptic responses were not altered (Supplementary Fig. 14a), E-LTP was impaired in neuropilin^{-/-} mice (Fig. 4d–f and Supplementary Fig. 14b, c; $P < 0.001$ versus wild-type at 20 min post-tetanus). These changes temporally correlated with neuropilin-mediated cleavage of EphB2, its dynamic interaction with NR1 and with the involvement of NMDA receptors in E-LTP in the lateral-basal pathway (Supplementary Fig. 15).

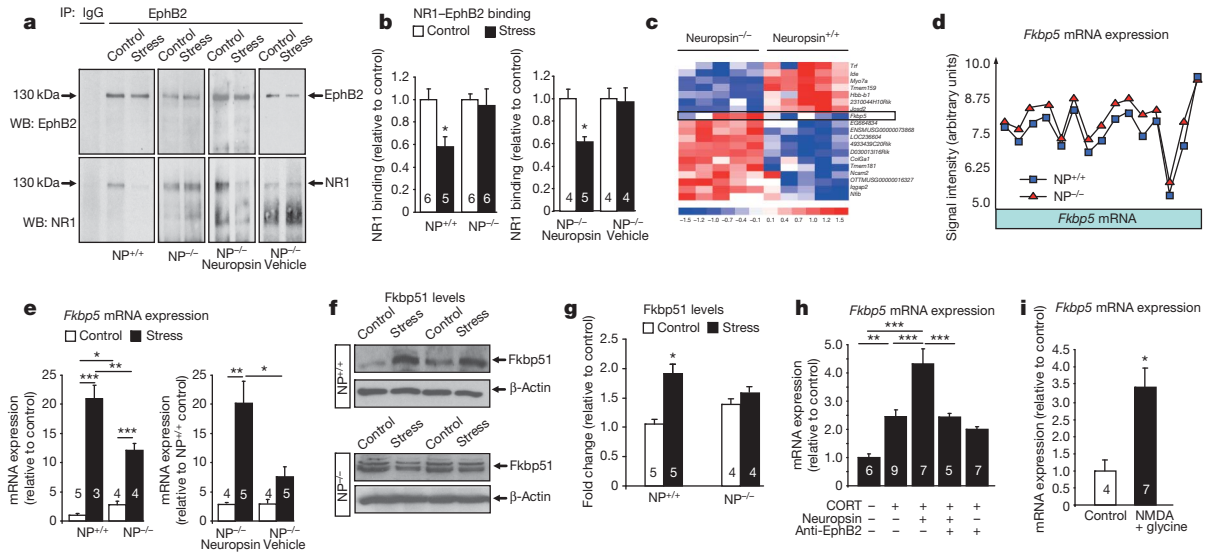


Figure 3 | Neuropsin regulates the dynamics of the EphB2–NR1 interaction and controls the expression of *Fkbp5*. **a, b**, EphB2 immunoprecipitation (IP; before and after 15 min of restraint stress) from amygdalae revealed dissociation of EphB2–NR1 complexes in wild-type ($F_{(3, 19)} = 4.2$; $P < 0.05$) but not neuropsin $^{-/-}$ mice. EphB2–NR1 dissociation was restored in neuropsin $^{-/-}$ mice by intra-amygdala neuropsin injections ($F_{(3, 13)} = 4.7$; $P < 0.05$). **c**, Microarray analysis of wild-type and neuropsin $^{-/-}$ amygdalae revealed differential expression of *Fkbp5* (heatmap in **c**, Supplementary Fig. 10). **d, e**, Exon-specific *Fkbp5* probes showed an upregulation of the whole transcript (**d**) confirmed by RT–qPCR (**e**; $F_{(3, 12)} = 72.15$; $P < 0.001$).

RT–qPCR revealed attenuated stress-induced upregulation of *Fkbp5* in neuropsin $^{-/-}$ mice (**e**; $P < 0.01$ for wild-type after stress versus neuropsin $^{-/-}$ after stress) rescued by intra-amygdala neuropsin injections ($F_{(3, 14)} = 9.2$; $P < 0.01$). **f, g**, *Fkbp51* protein levels were upregulated in wild-type mice ($F_{(3, 14)} = 8.95$; $P < 0.001$) but not in neuropsin $^{-/-}$ mice by stress. **h, i**, Neuropsin-mediated upregulation of *Fkbp5* in amygdala neuronal cultures ($F_{(4, 29)} = 19.04$; $P < 0.0001$) was blocked by anti-EphB2 antibody and mimicked by stimulation of NMDA receptors (**i**; $P < 0.05$). CORT, corticosterone. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Results are shown as mean \pm s.e.m. NP, neuropsin. Digits inside columns indicate n .

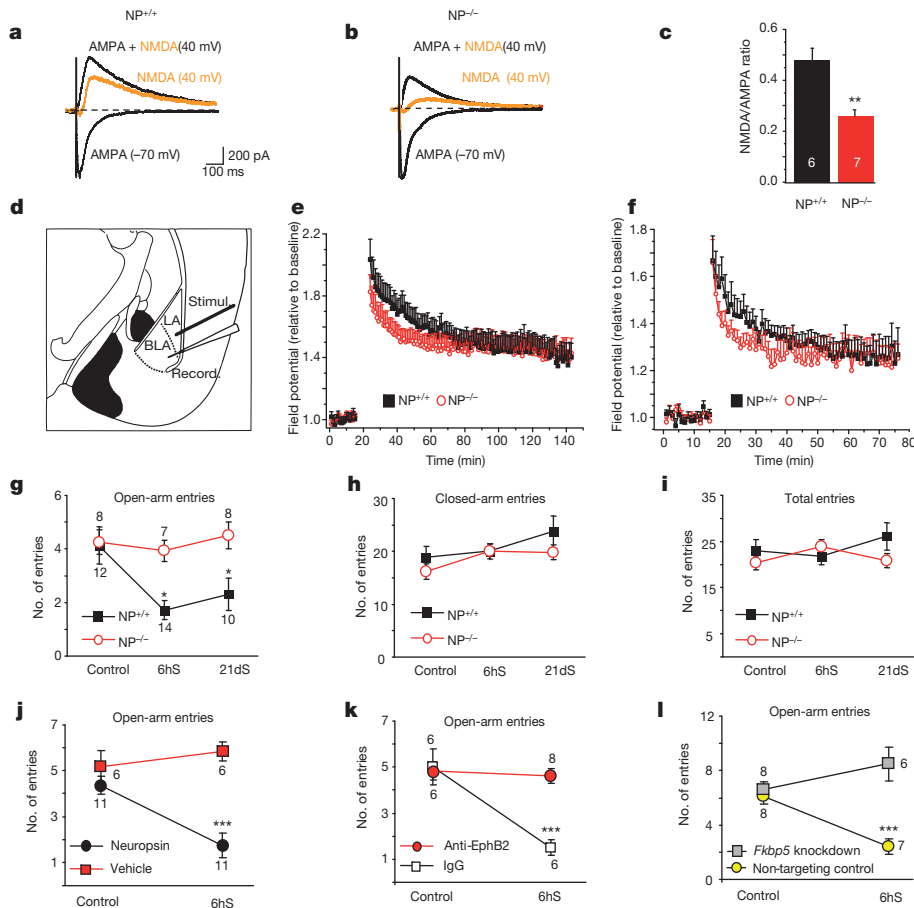


Figure 4 | Neuropsin controls NMDA receptor current, E-LTP and stress-induced anxiety. **a–c**, Whole-cell recordings from basal nucleus neurons of neuropsin $^{-/-}$ mice demonstrated lower NMDA currents compared to wild-type mice. **d–f**, Induction of LTP in the lateral-basal pathway (**d**) using a strong (**e**) or weak (**f**) protocol revealed an impairment of E-LTP in neuropsin $^{-/-}$ mice. BLA, basal amygdala; LA, lateral amygdala. Stimul., stimulation; Record., recording. **g**, The elevated-plus maze test after acute or chronic restraint stress demonstrated a lack of anxiety in neuropsin $^{-/-}$ mice as indicated by the number of entries into open arms. **h, i**, General locomotor activity was similar in both genotypes. **j**, The behavioural phenotype was reversed by bilaterally injecting neuropsin back into the amygdala of neuropsin $^{-/-}$ mice. **k, l**, Stress-induced anxiety in wild-type mice was disrupted by blocking EphB2 (**k**) or silencing the *Fkbp5* gene (**l**) in the amygdala. Digits inside columns or near symbols indicate n . 6hS, 6 h stress, 21dS, 21 days of daily restraint. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Results are shown as mean \pm s.e.m.

To examine if the neuropsin pathway alters behavioural signatures of stress we subjected wild-type and neuropsin^{-/-} mice to acute or chronic stress and measured anxiety in the elevated-plus maze (Fig. 4g–i). We found that stress caused a decrease in the number of entries of wild-type mice into open arms, indicative of high anxiety levels^{2,3}. In contrast, after stress, neuropsin^{-/-} mice did not develop anxiety (Fig. 4g; $P < 0.05$). Closed-arm entries, the total number of entries (Fig. 4h, i and Supplementary Fig. 16a, b) as well as general locomotor activity measures (Supplementary Fig. 17d) were similar between the genotypes as previously reported²³. Furthermore, neuropsin^{-/-} mice demonstrated an anxiolytic phenotype in the open field test, confirming a general role of neuropsin in regulating anxiety (Supplementary Fig. 17a–c). Although this effect is consistent with functional deficits in NMDA receptor function (Fig. 4a–c) and E-LTP (Fig. 4d–f) observed in neuropsin^{-/-} mice, it cannot be excluded that additional mechanisms, such as abnormal dendritic plasticity, may contribute to the lack of anxiety observed in neuropsin^{-/-} mice, particularly after long-lasting stress²⁴.

To examine whether the effect of neuropsin was acute and not associated with the lack of the protease during development we bilaterally injected neuropsin into the amygdalae of neuropsin^{-/-} mice (Supplementary Fig. 18). The neuropsin injection restored stress-induced anxiety in these animals (Fig. 4j; $P < 0.001$). The development of anxiety was hindered by blocking EphB2 in the amygdala of wild-type mice (Fig. 4k; $P < 0.001$), consistent with neuropsin interacting with EphB2 to facilitate stress-induced behavioural changes. Similarly, stress-induced anxiety was blocked by silencing *Fkbp5* gene expression in this brain region (Fig. 4l and Supplementary Fig. 19), consistent with a downstream role of *Fkbp5* in the neuropsin pathway.

Our studies favour a model where, after stress, both corticosterone-induced and neuropsin-mediated components converge to modulate *Fkbp5* gene expression and trigger anxiety (Supplementary Fig. 20). Neuropsin cleaves the extracellular portion of EphB2 and facilitates the dynamic interaction of EphB2 with the NR1 subunit of the NMDA receptor. The resulting enhancement of the NMDA current causes an upregulation of *Fkbp5* and promotes the development of anxiety. This novel pathway, highlighting the ability of Eph and NMDA receptors to respond to activity-dependent signals from the extracellular milieu, opens new possibilities for the treatment of stress-associated disorders, including various forms of anxiety disorders.

METHODS SUMMARY

Restraint stress was performed by placing the mice in wire mesh restrainers while control mice were left undisturbed. Anxiety was measured using the elevated-plus maze by counting the number of entries to closed or open arms during 5 min. Intra-amygdala injections were performed through bilaterally implanted cannulae and were followed by restraint stress in plexiglass tubes. The *Fkbp5* gene was silenced by intra-amygdala injection of lentiviral short hairpin RNA construct followed by behavioural assessment two weeks later. LTP was recorded from the lateral-basal pathway and whole-cell recordings made from basal amygdala neurons. Data were analysed by Student's *t*-test or ANOVA followed by Tukey's post-test. *P* values of less than 0.05 were considered significant.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 November 2010; accepted 15 February 2011.

Published online 20 April 2011.

- Gogolla, N., Caroni, P., Luthi, A. & Herry, C. Perineuronal nets protect fear memories from erasure. *Science* **325**, 1258–1261 (2009).
- Matys, T. *et al.* Tissue plasminogen activator promotes the effects of corticotropin-releasing factor on the amygdala and anxiety-like behavior. *Proc. Natl Acad. Sci. USA* **101**, 16345–16350 (2004).

- Pawlak, R., Magarinos, A. M., Melchor, J., McEwen, B. & Strickland, S. Tissue plasminogen activator in the amygdala is critical for stress-induced anxiety-like behavior. *Nature Neurosci.* **6**, 168–174 (2003).
- Pawlak, R. *et al.* Tissue plasminogen activator and plasminogen mediate stress-induced decline of neuronal and cognitive functions in the mouse hippocampus. *Proc. Natl Acad. Sci. USA* **102**, 18201–18206 (2005).
- Lupien, S. J., McEwen, B. S., Gunnar, M. R. & Heim, C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Rev. Neurosci.* **10**, 434–445 (2009).
- Roozendaal, B., McEwen, B. S. & Chattarji, S. Stress, memory and the amygdala. *Nature Rev. Neurosci.* **10**, 423–433 (2009).
- Lin, K. T., Slonowski, S., Ethell, D. W. & Ethell, I. M. Ephrin-B2-induced cleavage of EphB2 receptor is mediated by matrix metalloproteinases to trigger cell repulsion. *J. Biol. Chem.* **283**, 28969–28979 (2008).
- Klein, R. Bidirectional modulation of synaptic functions by Eph/ephrin signaling. *Nature Neurosci.* **12**, 15–20 (2009).
- Pasquale, E. B. Eph-ephrin bidirectional signaling in physiology and disease. *Cell* **133**, 38–52 (2008).
- Chen, Z. L. *et al.* Expression and activity-dependent changes of a novel limbic-serine protease gene in the hippocampus. *J. Neurosci.* **15**, 5088–5097 (1995).
- Bouvier, D. *et al.* Pre-synaptic and post-synaptic localization of EphA4 and EphB2 in adult mouse forebrain. *J. Neurochem.* **106**, 682–695 (2008).
- Perrotti, L. I. *et al.* Induction of Δ FosB in reward-related brain structures after chronic stress. *J. Neurosci.* **24**, 10594–10602 (2004).
- Matsumoto-Miyai, K. *et al.* NMDA-dependent proteolysis of presynaptic adhesion molecule L1 in the hippocampus by neuropsin. *J. Neurosci.* **23**, 7727–7736 (2003).
- Shimizu, C. *et al.* Characterization of recombinant and brain neuropsin, a plasticity-related serine protease. *J. Biol. Chem.* **273**, 11189–11196 (1998).
- Dalva, M. B. *et al.* EphB receptors interact with NMDA receptors and regulate excitatory synapse formation. *Cell* **103**, 945–956 (2000).
- Grunwald, I. C. *et al.* Kinase-independent requirement of EphB2 receptors in hippocampal synaptic plasticity. *Neuron* **32**, 1027–1040 (2001).
- Takasu, M. A., Dalva, M. B., Zigmond, R. E. & Greenberg, M. E. Modulation of NMDA receptor-dependent calcium influx and gene expression through EphB receptors. *Science* **295**, 491–495 (2002).
- Binder, E. B. The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology* **34**, S186–S195 (2009).
- Binder, E. B. *et al.* Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *J. Am. Med. Assoc.* **299**, 1291–1305 (2008).
- Binder, E. B. *et al.* Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nature Genet.* **36**, 1319–1325 (2004).
- Segman, R. H. *et al.* Peripheral blood mononuclear cell gene expression profiles identify emergent post-traumatic stress disorder among trauma survivors. *Mol. Psychiatry* **10**, 500–513 (2005).
- Genander, M. *et al.* Dissociation of EphB2 signaling pathways mediating progenitor cell proliferation and tumor suppression. *Cell* **139**, 679–692 (2009).
- Horii, Y., Yamasaki, N., Miyakawa, T. & Shiosaka, S. Increased anxiety-like behavior in neuropsin (kallikrein-related peptidase 8) gene-deficient mice. *Behav. Neurosci.* **122**, 498–504 (2008).
- Vyas, A., Mitra, R., Shankaranarayana Rao, B. S. & Chattarji, S. Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. *J. Neurosci.* **22**, 6810–6818 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by a Medical Research Council Project Grant (G0500231/73852), and a Marie Curie Excellence Grant (MEXT-CT-2006-042265 from the European Commission) to R. Pawlak, a Medisearch Fellowship to B.K.A., the Ministry of Education, Culture, Sports, Science and Technology of Japan (Grants-in-Aid, # 20300128) and Japan Science and Technology Agency (CREST program) to S.S., GENADDICT (LSHM-CT-2004-005166 from the European Commission) to R. Przewlocki, and NN405 274137 from the Ministry of Science and Higher Education of Poland to M.K. We are obliged to A. Kania, H. Castro and R. Fern for reagents. We are grateful to B. McEwen and A. Tobin for their comments on the manuscript and T. Gerdjikov, G. Mallucci and her laboratory members for advice.

Author Contributions B.K.A., J.-M.B., S.P., M.M., E.S., A.E.S., K.W.Y. and R. Pawlak performed experiments and data analysis; R. Pawlak and B.K.A. designed most of the experiments; S.S. provided neuropsin^{-/-} mice, suggestions and reagents; M.K., M.P. and R. Przewlocki performed bioinformatic analyses. R. Pawlak conceived the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Gene expression array data have been deposited in Gene Expression Omnibus under accession number GSE27088. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R. Pawlak (rp135@le.ac.uk).

METHODS

Mice. Experiments were performed on three-month-old wild-type (C57BL/6J) or neuropsin^{-/-} mice backcrossed to C57BL/6J for 12 generations. To generate neuropsin^{-/-} mice, exons 1–3 of the neuropsin gene, including the protease active site, were replaced by a neomycin resistance cassette²⁵. A lack of full-length neuropsin transcript and proteolytic activity in the brain of these mice was confirmed by RT–PCR and amidolytic assay²⁵, respectively. Neuropsin^{-/-} mice were genotyped as described¹³. Mice were housed three to five per cage in a colony room with a 12 h light/dark cycle (lights on at 07:00) with *ad libitum* access to commercial chow and tap water. The experiments were approved by the UK Home Office and the University of Leicester Ethical Committee.

Restraint stress. C57BL/6J and neuropsin^{-/-} mice were kept undisturbed for at least one week in their home cages and restraint stress was performed during the light period of the circadian cycle as described³. Control animals were left undisturbed, and stressed animals were subjected to a single 5 min, 15 min or 6 h restraint stress in a separate room. The mice were placed in their home cages in wire mesh restrainers secured at the head and tail ends with clips.

Primers. The primers for *Ephb2* (forward 5'-CTTCTCATCGCTGTGGTC and reverse 5'-ATGTGTCCGCTGGTGTAGTG) and *Fkbp5* (forward 5'-ATTGAT TGCCGAGATGTG and reverse 5'-TCTTACCAGGGCTTTGTC) were bought from Invitrogen. To quantify gene expression the target genes were compared against the actin gene as previously described²⁶.

RNA extraction. Control and stressed mice were anaesthetized (intraperitoneal sodium pentobarbital 50 mg kg⁻¹) and perfused transcardially (ice-cold PBS). Amygdalae were dissected from a coronal slice –0.58 to –2.3 mm relative to Bregma and stored in 'RNA later' (QIAGEN) at 4 °C. Alternatively, RNA was extracted from primary amygdala neuronal cultures using QIAzol lysis reagent (QIAGEN) and Mini Spin Columns according to the manufacturers' instructions (RNeasy Lipid tissue mini kit, QIAGEN). Two micrograms of RNA was converted to cDNA using Superscript III (Invitrogen) and oligo(dT) primers according to manufacturer's instructions.

RT–qPCR reaction. Triplicate wells contained 20 µl of SYBR Green Master Mix (Applied Biosystems and BioRad), forward primer (250 nM), reverse primer (250 nM), cDNA (1 µl) and nuclease free water to a total of 40 µl. PCR was performed using Chromo4/PTC-200 thermal cycler (MJ Research) under the following conditions: (1) 95 °C for 15 min; (2) 94 °C for 15 s; (3) 55 °C for 30 s; (4) 72 °C for 30 s; (5) steps 2–4 repeated 40 times. Control reactions were performed without DNA template and/or with unconverted RNA as the template.

Western blotting, cell fractionation and immunoprecipitation. Mice were anaesthetized (intraperitoneal sodium pentobarbital 50 mg kg⁻¹) and perfused transcardially (ice-cold PBS). The brains were removed and amygdalae dissected from a slice –0.58 to –2.3 mm relative to Bregma. Samples were homogenized in 0.1 M Tris, 0.1% Triton X-100, pH 7.4, containing phosphatase inhibitors (10 mM NaF, 1 mM Na orthovanadate) and protease inhibitors (Complete, Roche). Protein concentration was adjusted (Bradford method; Pierce). For Fkbp51 levels samples were homogenized in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, 10 mM NaF, 1 mM Na orthovanadate, complete Roche protease inhibitors, pH 8.0). Reduced (DTT) and denatured (100 °C for 5 min) samples (40 µg per lane) were subjected to SDS–PAGE electrophoresis and transferred onto nitrocellulose membrane, blocked (5% skim milk for 1 h at room temperature (25 °C)) and washed with TBS-T (3 × 5 min). The membranes were probed with the following primary antibodies overnight at 4 °C: goat anti-NCAM L1 (SantaCruz Biotechnology; 1:300), goat anti-EphB2, anti-EphB6 and anti-ephrinB2 (R&D; 1:500, 1:500 and 1:300, respectively), mouse anti-EphA4 (Zymed; 1:1,000), rabbit anti-pan-cadherin (Abcam; 1:2,000), rabbit anti-p75NGF (Chemicon, 1:1,000), rabbit anti-NR1 (Upstate, 1:250), rabbit anti-neuropsin (H. C. Castro), rabbit anti-Fkbp51 (Abcam; 1:250). The membranes were then washed in TBS-T (3 × 10 min) and incubated with a relevant HRP-conjugated secondary antibody as appropriate (Vector Labs, 1:1,000, 1 h, room temperature). The signal was developed, after washing with TBS-T (6 × 10 min), using a Western Blot Luminol Reagent (Santa Cruz). To normalize the results the membranes were stripped, blocked, washed as above and re-blotted using mouse anti-β-actin antibody (Sigma, 1:2,500, 1 h, room temperature). The membranes were then prepared and developed as above. To quantify the results the band intensities were measured using Scion Image software and normalized to the actin bands.

When indicated, cellular fractions of the amygdala samples were separated using a cellular protein fractionation kit (PerkinElmer) as per the manufacturer's protocol, and analysed by western blotting.

For immunoprecipitation, amygdala samples were homogenized as previously described²⁷, pre-cleared using goat IgG (Sigma, 1 µg) before incubation with goat anti-EphB2 antibody (R&D, 2 µg) for 1 h (4 °C). The samples were then incubated with protein G-sepharose beads overnight before being washed with PBS and analysed by western blotting.

Cell culture. SH-SY5Y cells at 80–90% confluence were washed with PBS three times before incubation with PBS, PBS and neuropsin (50 nM; R&D) or PBS and tissue plasminogen activator (tPA) (Alteplase, Genentech; 1 µg ml⁻¹) with or without human plasminogen for 15 min, after which the dishes were placed on ice and protease inhibitors (Complete, Roche) were added. The cells were collected and homogenized in Tris 50 mM pH 7.5, NaCl 150 mM, EDTA 5 mM, EGTA 5 mM, Triton-100 1%, NP40 0.5%. The resulting protein sample was analysed by western blotting as described above.

SH-SY5Y and HEK293 cells were transfected with mouse EphB2-GFP and incubated with PBS or PBS + neuropsin (300 nM) for 15 or 45 min. The samples were then treated as above for analysis.

For imaging, SH-SY5Y cells were transfected with GFP, mouse EphB2-GFP or EphA4-GFP (gift from A. Kania) and loaded with the cell tracker (Invitrogen). Images were taken with Zeiss LSM5 Exciter before and after 15 min incubation with neuropsin (50 nM, R&D), converted to greyscale and the intensity of the fluorescent signal was analysed using Scion Image.

Primary neuronal cultures were prepared from P1 C57BL/6J wild-type mice. The amygdalae were dissected and placed into a Petri dish (9.1 mM glucose, 25 mM HEPES, 5 mM KCl and 120 mM NaCl). Tissue was chopped and incubated in 10 ml of buffer containing 5 mg of protease (Type XIV; Sigma) and 5 mg of thermolysin (Type X; Sigma) at room temperature for 30–45 min. The digestion solution was replaced with 3 ml of HBSS (Gibco) plus 40 µg ml⁻¹ DNase. The mixture was titrated, centrifuged and resuspended in the plating medium (Neurobasal A medium, 10% fetal bovine serum, 2 mM Glutamax and 2% B-27 supplement, 100 µg ml⁻¹ streptomycin, 100 U ml⁻¹ penicillin (Invitrogen)), centrifuged and resuspended again. 20–30 µl droplets containing cells were added to the centre of poly-D-lysine (Sigma) coated coverslips and the plating medium was added one hour later. 5 µM cytosine B-D-arabinofuranoside (Ara-C; Sigma) was added to prevent proliferation of glial cells. Neurons were maintained in serum-free Neurobasal A medium at 37 °C in a humidified atmosphere of 5% CO₂/95% air. Half of the medium was replaced every 3–4 days. Cells were maintained for 11–16 days *in vitro* and then treated with either vehicle, corticosterone (10 nM), neuropsin (50 nM) or NMDA (100 µM, Sigma) + glycine (10 µM, Sigma). To block EphB2, neurons were treated with anti-EphB2 antibody (2 µg ml⁻¹; R&D) 10 min before the experiment.

Electrophysiology. For field recordings coronal slices of the amygdala (400 µm) were obtained from 8–12 weeks-old neuropsin^{-/-} and wild-type mice. The animals were anaesthetized with ketamine/xylazine (2:1 ratio; 2.4 µl g⁻¹ i.p.). Slices were prepared using a vibrating microtome (Campden Instruments; MA752) in ice-cold, low sodium ACSF (sucrose 249 mM, KCl 2.5 mM, NaH₂PO₄ 1.25 mM, D-glucose 10 mM, NaHCO₃ 26 mM, CaCl₂ 0.1 mM, MgSO₄ 2.9 mM, ascorbic acid 0.5 mM, bubbled with 95% O₂/5% CO₂ mixture, pH 7.3). Slices were placed in a holding chamber for 30 min at 35 °C and then for at least 2.5 h (30 min for whole-cell recordings) at room temperature (25 °C) in ACSF (NaCl 124 mM, KCl 5 mM, NaH₂PO₄ 1.25 mM, D-glucose 10 mM, NaHCO₃ 26 mM, CaCl₂ 2.4 mM, MgSO₄ 1.3 mM). All the experiments were performed at room temperature.

Extracellular recording were made with a bipolar tungsten electrode (WPI). For recordings, glass microelectrodes (1–2 MΩ) filled with ACSF were used. To record field potentials in the lateral-basal amygdala pathway the stimulating electrode was positioned in the lateral amygdaloid nucleus close to the external capsule and the recording electrode in the basal nucleus²⁸. The stimulus intensity was adjusted to evoke a field potential (FP) of 60–70% (0.2 ms pulse duration) of the maximal amplitude. The amygdala was stimulated every 30 s in order to record a stable baseline for at least 15 min. Several responses were averaged and a template was created. Only the responses matching the template were analysed. E-LTP was evoked by a single tetanic stimulation (100 Hz, 1 s). L-LTP were elicited by two tetanic stimulations (100 Hz, 1 s, 10 s interval) repeated 4 times in 3 min intervals with the same intensity and pulse duration as the test stimuli as described²⁹. The recordings were amplified (Multiclamp 700b, Axon Instruments), filtered (10 kHz) and digitized at 50 kHz (Digidata 1440A, Axon Instruments). pClamp 10 (Axon Instruments) and Origin 7 (Microcal) software were routinely used during data acquisition and analysis.

For whole-cell recordings, coronal slices (300 µm) were obtained from 3–4-week-old mice. The animals were anaesthetized with hypnorm/midazolam (1:1; 8 µl g⁻¹ body weight, intraperitoneally). Recordings were made from somata of principal neurons of the basal nucleus of the amygdala. Principal neurons and interneurons were distinguished by their morphological and electrophysiological properties³⁰. After whole-cell configuration the series resistance was regularly monitored and a maximum of 10–15 MΩ tolerated. AMPA and NMDA currents were recorded by clamping the membrane potential of the cell at –70 mV and +40 mV respectively (average of 5 traces each). The slice was subsequently perfused with AP-5 in order to isolate the NMDA component at +40 mV (subtraction of traces before and after perfusion with AP-5). The NMDA/AMPA ratio was

obtained by measuring the peak of the AMPA current at -70 mV in the presence of AP-5 and the peak of NMDA current at $+40$ mV. At the end of the experiments the currents were blocked by CNQX ($30 \mu\text{M}$ in DMSO) and AP-5 ($50 \mu\text{M}$ in DMSO). All drugs were bath applied (perfusion rate 1.5 ml min^{-1}). The recording electrodes were borosilicate glass pipettes ($2-4 \text{ M}\Omega$). The pipettes were filled with the following solution: Cs-methyl sulphonate 130 mM , KCl 8 mM , EGTA 0.5 mM , HEPES 10 mM , glucose 5 mM , QX314 5 mM . ACSF composition: NaCl 124 mM , KCl 5 mM , NaH_2PO_4 1.25 mM , D-glucose 10 mM , NaHCO_3 26 mM , CaCl_2 2 mM , MgSO_4 1 mM . All the experiments were performed at 25°C . Data were recorded with a Multiclamp 700B amplifier, filtered at 10 kHz and digitized at 50 kHz (Digidata 1440A, Axon instruments). pClamp 10 (Axon Instruments) and Origin 7 (Microcal) software was routinely used during data acquisition and analysis.

Analysis of the Fkbp5 promoter. The identification of over-represented transcription factor binding sites (TFBSs) in the promoter region of *Fkbp5* was performed using the cREMaG database (<http://cremag.org>). The promoter region of *Fkbp5* was defined as an evolutionarily conserved (on the basis of the alignment of mouse and human genes) sequence between $10,000 \text{ bp}$ upstream and $5,000 \text{ bp}$ downstream of the transcription start site (TSS). The parameters of 65% conservation threshold and maximum number of top $10,000$ conserved TFBSs in coding and non-coding regions were used. The obtained results were compared to conserved promoter background. Over-representation was measured as the number of identified TFBSs compared to putative number obtained by chance ($P < 0.01$).

Microarray study. Amygdalae were isolated from wild-type ($n = 15$) and neuropsin $^{-/-}$ ($n = 15$) mice using a dissecting microscope in ice-cold ACSF (glucose 25 mM , NaCl 115 mM , $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ 1.2 mM , KCl 3.3 mM , CaCl_2 2 mM , MgSO_4 1 mM , NaHCO_3 25.5 mM , pH 7.4 and stored at -20°C in RNAlater solution (Qiagen). RNA was extracted using RNeasy Lipid Tissue Mini Kit (Qiagen), the ribosomal fraction of RNA reduced with RiboMinus Kit (Invitrogen) and the RNA integrity verified by electrophoresis using Agilent Bioanalyser 2100 (Agilent Technologies). RNA pulled from three mice was reverse-transcribed and hybridized with GeneChip Mouse Exon 1.0 ST Array (Affymetrix; 5 arrays per genotype).

Microarray data were initially processed using GeneChip Operating Software. DTT data were transferred by Transfer Tool software (Affymetrix). Chip quality and raw microarray data pre-processing were performed according to the Affymetrix guidelines using Expression Console software (Affymetrix). After background subtraction, the data were processed using the RMA method and quantile normalization. The obtained results were taken as the measure of mRNA abundance derived from the level of gene expression. Significance levels (P values) of differences in mRNA abundance between the wild-type and neuropsin $^{-/-}$ animals were calculated for each probe set using the Student's t -test. The P values for all exons of each particular gene were multiplied to establish gene P value. The threshold of $P < 0.05$ for each gene was computed using a permutation test followed by Bonferroni correction for multiple testing. All the statistical analyses were done in R software version 2.8.1 (<http://www.r-project.org>). Sources of variation were analysed by a three-way ANOVA using Partek Genomic Suite.

Immunohistochemistry. Mice were anaesthetized (intraperitoneal sodium pentobarbital 50 mg kg^{-1}) and transcardially perfused (ice-cold PBS containing protease inhibitors (Complete, Roche) followed by ice-cold 4% paraformaldehyde). The brains were dissected and fixed in 4% paraformaldehyde in PBS overnight at 4°C . Seventy-micrometre-thick coronal slices were collected on a vibrating microtome and stored at 4°C in PBS containing 0.002% sodium azide (Sigma). Slices were preincubated in PBS-T (PBS solution 0.5% bovine serum albumin, 0.02% Triton X-100 and blocking sera at 1:500) for 5 h at room temperature, incubated with goat anti-EphB2 (1:300, R&D) or rabbit anti-neuropsin (1:200, H. Castro; the antibody was preabsorbed on acetone powder prepared from neuropsin $^{-/-}$ brain for 1 h at room temperature before use) antibodies, along with mouse anti-NeuN (1:200, Chemicon) and chicken anti-GFAP (1:1,000, Dako) overnight at 4°C in PBS-T. Next, the slices were washed for 8–10 h with PBS-T and incubated overnight with compatible FITC, Alexa Fluor 488, Alexa Fluor 546 or Alexa Fluor 647 secondary antibodies (1:500, Molecular Probes) in the same buffer. Control sections were processed with the primary antibodies omitted. For double Ephb2/neuropsin co-labelling sections were incubated in PBS-T containing anti-goat Alexa Fluor 546 as well as anti-rabbit Alexa Fluor 488 for detection of the above primary antibodies. TOTO-3 iodide (1 nM , Molecular Probes) was added to the secondary antibody mixture. For the triple Ephb2/neuropsin/Fkbp5 labelling rat anti-Fkbp51 (R&D, 1:500) was additionally used along with a compatible Alexa Fluor 647 secondary antibody. Sections were then washed in PBS-T for 5 h, mounted on glass slides using Vectamount medium (Vector Laboratories), and photographed using Zeiss LSM 5 Exciter confocal microscope.

Corticosterone levels. Mice were subjected to restraint stress of various durations with or without recovery and trunk blood was collected to measure corticosterone levels in the plasma by EIA according to the manufacturer's instructions (Cayman Chemicals, Cat No. 500651).

Elevated-plus maze. The elevated-plus maze test was performed as previously described³. The apparatus consisted of four non-transparent white Plexiglas arms: two enclosed arms ($50 \times 10 \times 30 \text{ cm}$) that formed a cross shape with the two open arms ($50 \times 10 \text{ cm}$) opposite each other. The maze was 55 cm above the floor and dimly illuminated. Wild-type and neuropsin $^{-/-}$ mice were tested 12 h after the restraint stress. Mice were placed individually on the central platform, facing an open arm, and allowed to explore the apparatus for 5 min. Behaviour was recorded by an overhead camera. The number of entries of the animal from the central platform to closed or open arms was counted. The maze was cleaned with 70% alcohol after each session to avoid any odorant cues.

Open field. Mice were placed in a $50 \times 50 \times 50 \text{ cm}$ plexiglas box and were left free to move during 10 min. The box was cleaned with 70% alcohol after each session to avoid any odorant cues. An overhead camera placed above the box recorded the session. Locomotor parameters were analysed with the ANY-MAZE software (Stoelting).

Stereotaxic injections. Mice were intraperitoneally anaesthetized with ketamine/xylazine (100 and 10 mg kg^{-1} , respectively), placed in a stereotaxic apparatus and bilaterally implanted with stainless steel guide cannulae (26 gauge; Plastics One, Roanoke, VA) aimed above the basolateral complex of the amygdala (1.5 mm posterior to Bregma, 3.5 lateral and 4.0 ventral). The cannulae were secured in place with dental cement. Dummy cannulae were inserted into all implanted cannulae to maintain patency. After one week dummy cannulae were replaced with the injection cannulae (projecting 0.75 mm from the tip of the guide cannulae to reach the basolateral complex of the amygdala) and the mice were injected with either anti-EphB2 antibody (R&D, $1 \mu\text{l}$, $2 \mu\text{g ml}^{-1}$), control IgG, recombinant neuropsin (R&D, $1 \mu\text{l}$, 50 nM) or its vehicle followed by 6 h restraint stress in transparent plexiglass tubes. After the experiment a small volume of bromophenol blue was injected to visualize the injection site, the brains were sectioned and the cannulae placement was determined histologically.

Fkbp5 gene silencing. To silence the *Fkbp5* gene in the amygdalae we used SMARTvector 2.0 lentiviral shRNA technology (Dharmacon) using a human cytomegalovirus (hCMV) promoter and a turboGFP reporter gene. Three different targeting constructs were tested. First, $0.3 \mu\text{l}$ of the lentivirus was injected at a point 1.7 mm posterior to Bregma, 3.5 mm lateral from the midline and 4.4 mm ventral at 200 nl min^{-1} using the Nanofil syringe with a 33-gauge needle through an UMP-3.1 micropump (all from World Precision Instruments) mounted on Stoelting stereotaxic frame. After 5 min the needle was lowered to 5 mm ventral and $0.3 \mu\text{l}$ of the virus injected. The needle remained in place for another 5 min to prevent the backflow, slowly removed and the skin closed with Vetbond (3M). After two-weeks recovery the amygdalae were dissected to determine the knock-down efficiencies *in vivo* as compared to the non-targeting construct (TGGTTTACATGTTGTGTGA; $2.66 \times 10^8 \text{ TU ml}^{-1}$) or uninjected mouse amygdalae by RT-qPCR and western blotting as described above. The most efficient construct ($\sim 60\%$ mRNA and protein knockdown efficiency; targeting sequence ATGCTGAGCTTATGTACGA; $3.02 \times 10^8 \text{ TU ml}^{-1}$) was used to silence the *Fkbp5* gene in all subsequent behavioural experiments. The restraint stress and the elevated-plus maze were performed two weeks after the intra-amygdala lentivirus injection as described above. The region specificity of lentiviral injections was verified histologically by direct observation of the turboGFP fluorescence on consecutive coronal sections spanning the amygdala, using a Zeiss LSM5 Exciter confocal microscope.

Statistics. Student's t -test (when two groups were compared) or analysis of variance (ANOVA) followed by Tukey's post-test were used as appropriate. P values of less than 0.05 were considered significant. The ANOVA P values are reported in the text, the results of the post-test are indicated by asterisks on graphs.

25. Hirata, A. *et al.* Abnormalities of synapses and neurons in the hippocampus of neuropsin-deficient mice. *Mol. Cell. Neurosci.* **17**, 600–610 (2001).
26. Salter, M. G. & Fern, R. NMDA receptors are expressed in developing oligodendrocyte processes and mediate injury. *Nature* **438**, 1167–1171 (2005).
27. Calo, L. *et al.* Interactions between ephrin-B and metabotropic glutamate 1 receptors in brain tissue and cultured neurons. *J. Neurosci.* **25**, 2245–2254 (2005).
28. Brambilla, R. *et al.* A role for the Ras signalling pathway in synaptic transmission and long-term memory. *Nature* **390**, 281–286 (1997).
29. Huang, Y. Y. & Kandel, E. R. 5-Hydroxytryptamine induces a protein kinase A/mitogen-activated protein kinase-mediated and macromolecular synthesis-dependent late phase of long-term potentiation in the amygdala. *J. Neurosci.* **27**, 3111–3119 (2007).
30. Washburn, M. S. & Moises, H. C. Electrophysiological and morphological properties of rat basolateral amygdaloid neurons *in vitro*. *J. Neurosci.* **12**, 4066–4079 (1992).

Stem-cell-triggered immunity through CLV3p–FLS2 signalling

Horim Lee¹, Ok-Kyong Chah¹ & Jen Sheen¹

Stem cells in the shoot apical meristem (SAM) of plants are the self-renewable reservoir for leaf, stem and flower organogenesis^{1,2}. In nature, disease-free plants can be regenerated from SAM despite infections elsewhere, which underlies a horticultural practice for decades³. However, the molecular basis of the SAM immunity remains unclear. Here we show that the CLAVATA3 peptide (CLV3p), expressed and secreted from stem cells and functioning as a key regulator of stem-cell homeostasis in the SAM of *Arabidopsis*^{1,2,4}, can trigger immune signalling and pathogen resistance via the flagellin receptor kinase FLS2 (refs 5, 6). CLV3p–FLS2 signalling acts independently from the stem-cell signalling pathway mediated through CLV1 and CLV2 receptors^{1,2,4}, and is uncoupled from FLS2-mediated growth suppression^{5,6}. Endogenous CLV3p perception in the SAM by a pattern recognition receptor for bacterial flagellin, FLS2, breaks the previously defined self and non-self discrimination in innate immunity^{6,7}. The dual perception of CLV3p illustrates co-evolution of plant peptide and receptor kinase signalling for both development and immunity. The enhanced immunity in SAM or germ lines may represent a common strategy towards immortal fate in plants and animals^{1,2,8}.

In both plants and animals, innate immunity is triggered through pattern recognition receptors (PRRs) in response to microbe-associated molecular patterns (MAMPs)^{6,7} to provide the first line of inducible defence. Plant receptor kinases represent the main functions of known plasma membrane PRRs for MAMP recognition to distinguish non-self from self⁶. FLS2 is the first characterized plant leucine-rich-repeat (LRR) receptor kinase that perceives bacterial flagellin and launches convergent downstream signalling and defence pathways for potentially broad-spectrum pathogen resistance^{5,6,9–11}. The perception of bacterial flagellin is conserved in seed plants, and functional FLS2 orthologues are found from *A. thaliana* to rice^{5,6}. As FLS2 is expressed throughout the whole plant including the SAM (Supplementary Fig. 2)^{5,12}, flagellin–FLS2 signalling could provide immune protection in different parts of the plant body after infections.

While developing a plant expression system to screen for peptide-mediated receptor-like kinase (RLK) signalling, we observed that the endogenously modified 12-amino-acid CLV3p (Supplementary Table 1)¹³ triggered similar responses as flg22 (the conserved 22-amino-acid peptide of bacterial flagellin) in mesophyll protoplasts^{5,6,9,14–16}. This finding was unexpected because CLV3p is normally expressed, secreted and processed by the stem cells to control SAM maintenance via CLV1 and CLV2 receptors^{1,2,4,13,17}. Flg22 and CLV3p, but not Δ CLV3p lacking the last His residue, activated similar mitogen-activated protein kinase (MAPK) activities detected by the in-gel kinase assay (Fig. 1a and Supplementary Fig. 3)^{6,9,14–16}. Highly purified CLV3p synthesized by different sources displayed the same activities, ruling out the contamination possibility. We sought to identify the CLV3p receptor in leaf cells by examining MAPK activation in various receptor mutants. Neither the dominant *clv1-1* mutant (Fig. 1a) nor the *clv2-1* mutant (Supplementary Fig. 4)^{1,2,4,17–19} affected CLV3p-triggered MAPK activation. Surprisingly, two independent *fls2* mutant

alleles of Landberg *erecta* (*Ler*) and Columbia (*Col-0*)^{5,6}, but not the *efr-1* (the bacterial elongation factor EF-Tu receptor EFR) mutant²⁰, failed to support the activation of MAPKs by both flg22 and CLV3p (Fig. 1a and Supplementary Fig. 3). Complementation with the wild-type *FLS2* gene (Fig. 1b), but not *CLV1* (data not shown) in the *fls2* mutant, confirmed that FLS2 could recognize both flg22 and CLV3p to mediate MAPK signalling. Importantly, CLV3p and flg22 activated similar early marker genes, including *FRK1*, *WRKY29* and *WRKY30* (Fig. 1c)^{9,10,14}.

FLS2 signalling requires the recruitment of the RLK BAK1 and the interaction between FLS2 and BAK1 represents the earliest event (within a minute) triggered by flg22 binding to FLS2 (refs 6, 14–16). Like flg22, CLV3p also induced the immediate interaction between FLS2 and BAK1 detected by reciprocal co-immunoprecipitation (Supplementary Fig. 5a, b)^{14–16}. Consistently, CLV3p signalling monitored by MAPK activation was greatly diminished in the *bak1-4* mutant (Supplementary Fig. 6)^{14–16}. Notably, CLV3p pre-treatment could confer enhanced resistance to the pathogenic bacteria *Pseudomonas syringae* pv. *tomato* DC3000 in an FLS2-dependent manner (Fig. 1d)^{10,14}. These comprehensive analyses strongly support our new finding that the conserved flagellin receptor FLS2 can recognize the stem-cell peptide CLV3p and trigger convergent innate immune signalling^{6,9,14}.

Peptide titration experiments using the FLS2 and BAK1 co-immunoprecipitation assay showed that 1 nM flg22 was as potent as 1 μ M CLV3p (Fig. 2a and Supplementary Fig. 5c) required for SAM suppression^{13,21}. Intriguingly, flg22 and CLV3p peptides supported similar primary gene activation but distinct long-term growth effects

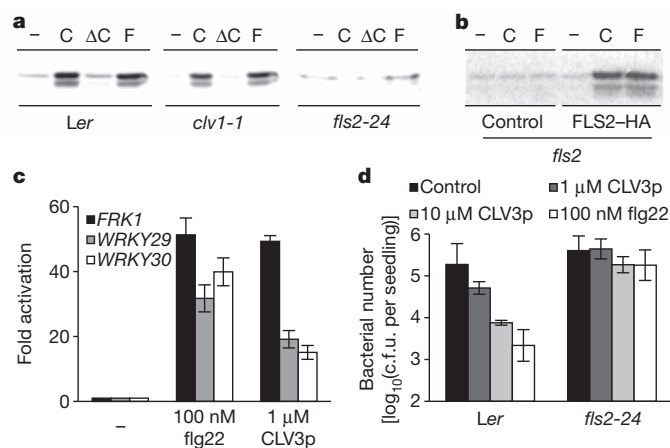


Figure 1 | CLV3p and flg22 activate similar downstream responses through FLS2. **a**, MAPK activity analysis. *Ler*, *clv1-1* and *fls2-24* protoplasts were treated with 1 μM CLV3p (C), 1 μM Δ CLV3p (Δ C) or 100 nM flg22 (F) for 10 min. **b**, FLS2–HA complements *fls2*. **c**, CLV3p triggers flg22 marker gene activation. Quantification by qRT–PCR; peptide treatment for 1 h. Error bars indicate s.d. ($n = 3$). **d**, CLV3p enhances resistance to *P. syringae* pv. *tomato* DC3000. Error bars indicate s.d. ($n = 3$). c.f.u., colony-forming units.

¹Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114, USA.

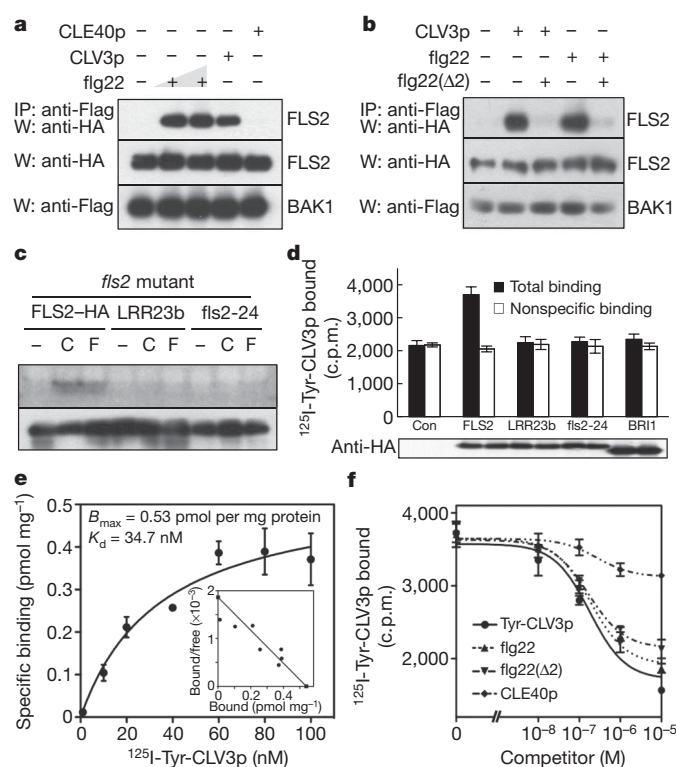


Figure 2 | CLV3p and flg22 share similar perception through FLS2.

a, CLE40p fails to induce FLS2–BAK1 interaction. Treatments: flg22 (1, 10 nM), 1 μ M CLV3p or 1 μ M CLE40p for 10 min. **b**, Flg22(Δ 2) blocks CLV3p and flg22 signalling. Flg22(Δ 2) (50 μ M). **c**, LRR mutations of FLS2 eliminate flg22 and CLV3p perceptions in *fls2*. Top, in-gel kinase assay; bottom, protein expression. **d**, 125 I-Tyr-CLV3p specifically binds to FLS2. Nonspecific binding (20 μ M unlabelled Tyr-CLV3p). Error bars indicate s.d. ($n = 3$). c.p.m., counts per min. **e**, Saturation binding curve and Scatchard plot. Error bars indicate s.d. ($n = 3$). **f**, Binding competition analysis with peptides. Error bars indicate s.d. ($n = 3$). Nonspecific binding (10 μ M unlabelled Tyr-CLV3p).

in seedling assays (Figs 3 and 4 and Supplementary Fig. 7). Flg22 is a MAMP from non-self invaders and is detected by a very sensitive perception system²² to induce innate immunity in a timely manner after infection. CLV3p, on the other hand, is an endogenous plant signal naturally secreted in the stem-cell zone to activate constitutive innate immunity via FLS2 in the SAM, which might provide a type of ‘vaccination’ before any infections to elicit sufficient immune protection without severe growth penalty caused by MAMPs (Fig. 3a–e)^{5,6,15,20}.

CLV3 belongs to the CLV3/ERS-related (CLE) gene family, which is conserved in diverse plant species^{4,17–19}. The *Arabidopsis* genome has 32 CLE genes^{4,17–19}. CLV3p and many other CLE peptides share similar root growth inhibition activity through the CLV2–CRN (CORYNE)/SOL2 (SUPPRESSOR OF OVEREXPRESSION OF LLP1-2) receptor complex, revealing a high degree of redundancy in peptide signalling^{18,19}. To assess the specificity of CLV3p–FLS2 signalling, we examined the activity of synthetic peptide CLE40p, which belongs to the same CLE subgroup as CLV3p (refs 4, 17, 18, 21). CLE40p triggered the growth arrest of SAM, repression of *WUSCHEL* (*WUS*) detected by *pWUS::GUS* (ref. 23), and root growth inhibition (Supplementary Fig. 8a, b) as previously reported for CLV3p and CLE19p (refs 4, 17, 18, 21). *WUS* encodes a homeodomain transcription factor that has a central role in *Arabidopsis* SAM maintenance¹². However, CLE40p did not promote FLS2–BAK1 interaction or flg22 early marker gene activation (Fig. 2a and Supplementary Fig. 8c). CLV3p–FLS2 signalling might be unique among CLE peptides.

To determine whether flg22- and CLV3p-triggered FLS2 signalling through the same or different extracellular sites, we carried out competition experiments using a well-established antagonist peptide lacking only

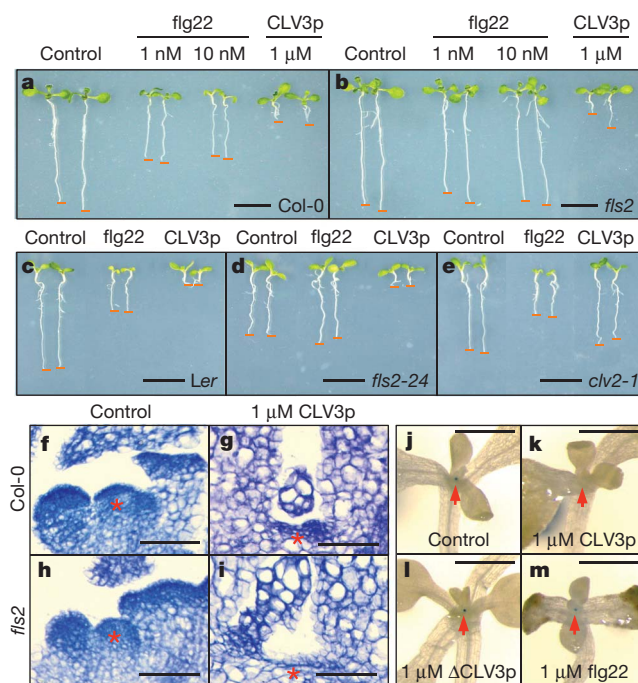


Figure 3 | CLV3p-mediated SAM arrest and immune signalling are uncoupled from flg22-triggered growth suppression. **a–e**, Growth inhibition analysis. Treatment with flg22 (1, 10 nM) or CLV3p (1 μ M) in Col-0 (**a**), *fls2* (**b**), *Ler* (**c**), *fls2-24* (**d**) and *clv2-1* (**e**) seedlings. Scale bars, 1 cm. **f–i**, CLV3p suppresses the SAM in Col-0 and *fls2* plants. Red asterisks indicate SAM region. Scale bars, 50 μ m (**f–i**). **j–m**, CLV3p represses *pWUS::GUS* expression. Treatment was without (**j**) or with 1 μ M CLV3p (**k**), 1 μ M Δ CLV3p (**l**) or 1 μ M flg22 (**m**). Red arrows indicate SAM region. Scale bars, 1 mm.

the last two amino acids at the carboxy terminus of flg22 (flg22(Δ 2))²⁴. On the basis of the CLV3p- or flg22-triggered FLS2–BAK1 interaction assay by co-immunoprecipitation, we showed that flg22(Δ 2) effectively competed with and blocked flg22 and CLV3p signalling (Fig. 2b). Moreover, two LRR mutants, one in the 10th LRR domain (*fls2-24*) and the other in the 23rd LRR domain (the *LRR23b* mutant) of FLS2, failed to activate flg22 and CLV3p signalling detected by MAPK activation despite normal FLS2 protein levels (Fig. 2c)^{9,14–16,24–26}. These results indicate that flg22 and CLV3p probably share binding and activation sites in the extracellular LRR domain of FLS2.

On the basis of the established studies on the flg22–FLS2 and CLV3p–CLV1 interactions^{27,28}, we developed a cell-based assay to show that 125 I-Tyr-CLV3p interacted directly with FLS2 expressed in the null *fls2* mutant protoplasts. Importantly, only the wild-type FLS2 protein but not the *fls2-24* and *LRR23b* mutant protein or the brassinosteroid receptor kinase (BRI1) showed specific binding to 125 I-Tyr-CLV3p (Fig. 2d). The saturation binding curve and Scatchard plot were generated using the specific binding assay. The estimated dissociation constant (K_d) for FLS2 and 125 I-Tyr-CLV3p interaction was 34.7 nM (Fig. 2e), which was close to the K_d for CLV1 and CLV3p interaction²⁸. The specific binding could be competed by the unlabelled Tyr-CLV3p, CLV3p, flg22, flg22(Δ 2), but not CLE40 (Fig. 2f and Supplementary Fig. 9). Tyr-CLV3p and CLV3p displayed identical effectiveness in competing with 125 I-Tyr-CLV3p for binding to FLS2 (Supplementary Fig. 9), immune marker gene activation and root inhibition (data not shown). 125 I-Tyr-CLV3p binding to FLS2 shared characteristics of CLV3p binding to CLV1 (ref. 28). Although the estimated K_d for 125 I-Tyr-flg22 binding to FLS2 is lower²⁷ and FLS2 responses to flg22 are more sensitive (Figs 1c, d and 2a, and Supplementary Fig. 5a, c), flg22 did not compete more effectively than Tyr-CLV3p and CLV3p for 125 I-Tyr-CLV3p binding to FLS2. Because the sequences of CLV3p and flg22 do not share any overt similarity and Δ CLV3p was ineffective in activating or blocking FLS2 signalling

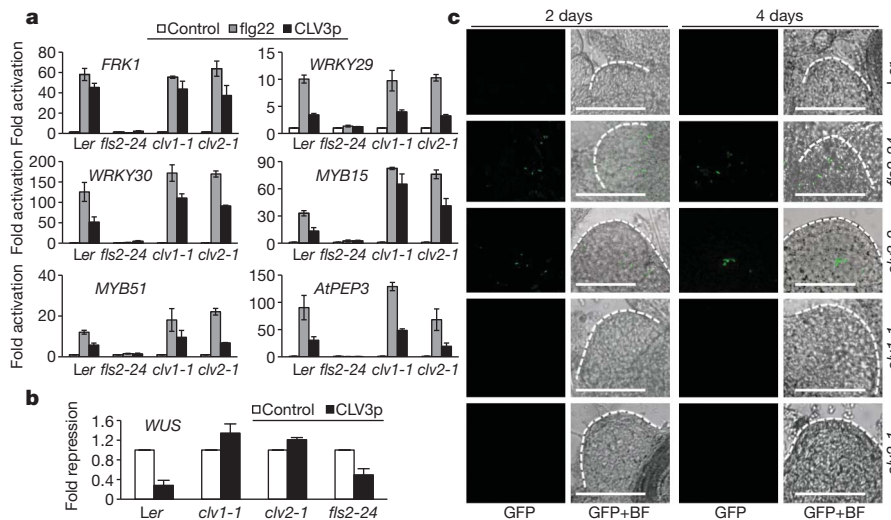


Figure 4 | CLV3p-FLS2 signalling enhances innate immunity for SAM protection. **a**, Flg22 marker gene activation by CLV3p through FLS2 in the SAM. **b**, *WUS* repression by CLV3p was mediated via CLV1 or CLV2 but not FLS2. The SAM tissues were analysed by qRT-PCR after treatment with 1 nM flg22 or 1 μ M CLV3p for 1 h. Error bars indicate s.d. ($n = 3$). **c**, Infection and

(Fig. 1a, data not shown), it was surprising to discover that they could still compete for some potentially shared binding sites on the LRR of FLS2 (Fig. 2b, c, f). The precise location of flg22 and CLV3p binding to FLS2 awaits future co-crystallographic analysis or mass spectrometry studies on cross-linked ligand-receptor complexes.

In previous studies, innate immune responses and growth inhibition are always tightly linked^{5,6,15,20}. We investigated the effects of flg22 and CLV3p on seedling growth^{5,6,15,20}. Surprisingly, CLV3p did not inhibit shoot growth but caused stronger root growth arrest in Col-0 and *Ler* wild-type seedlings (Fig. 3a, c and Supplementary Figs 7a and 8b). Whereas seedling growth inhibition in both shoots and roots by flg22 was eliminated in two independent *fls2* mutants, the stronger root growth arrest by CLV3p was retained (Fig. 3b, d). Although CLV3p activated, via FLS2, a spectrum of innate immune responses similar to those via flg22-FLS2 signalling in mesophyll cells, seedlings and the SAM (Figs 1, 2 and 4 and Supplementary Figs 3, 5 and 8), CLV3p did not stimulate the typical flg22-FLS2-mediated growth suppression in whole seedlings (Fig. 3a–e). Similar to other synthetic CLE peptides, CLV3p-triggered root growth arrest was abolished in the *clv2-1* mutant (Fig. 3c–e and Supplementary Fig. 7)^{18,19}. Thus, CLV3p-FLS2 signalling activated only immune responses but not the general growth inhibition. Significantly, CLV3p was very active in triggering the typical SAM arrest in null *fls2* seedlings (Fig. 3f–j)^{1,2,4,13,17–19,23,28}. The SAM growth arrest was correlated with the repression of a sensitive *pWUS::GUS* reporter in the organizing centre by the exogenous CLV3p but not Δ CLV3p or flg22 (Fig. 3j–m)²³.

Because CLV3p is specifically expressed and secreted from stem cells of the SAM^{1,2,4,13,17,23}, it is critical to examine CLV3p-FLS2 signalling in the SAM to evaluate its physiological relevance. Using quantitative real-time reverse transcriptase PCR (qRT-PCR) analysis with isolated SAM tissues, CLV3p clearly triggered two parallel signalling pathways in the SAM (Fig. 4a, b and Supplementary Figs 10–13). The activation of important immune marker genes by CLV3p in wild type, *clv1-1*, *clv2-1* but not *fls2-24* validated the action of innate immune signalling via FLS2 in the *Arabidopsis* SAM. These flg22 and CLV3p inducible genes, including *FRK1* (a RLK), *WRKY29*, *WRKY30*, *MYB15*, *MYB51* (transcription factors) and *PEP3* (a peptide inhibiting *P. syringae* pv. *tomato* DC3000 growth via RLK signalling), have important roles in bacterial and fungal resistance^{6,9,10,14}. Consistently, some of these marker genes showed reduced endogenous expression in the SAM of *clv3-2*, but were constitutively expressed at higher levels in the SAM but not other tissues of wild type (Supplementary Figs 10, 11a and

proliferation of pathogenic bacteria in the SAM lacking CLV3p-FLS2 signalling. *P. syringae* pv. *tomato* DC3000–GFP was co-cultivated with seedlings for 2 or 4 days. *P. syringae* pv. *tomato* DC3000–GFP was visualized using a confocal microscope. Scale bars, 50 μ m. BF, bright field.

12). Their endogenous expression levels were low in both the SAM and other tissues in *clv3-2* and *fls2-24* (Supplementary Fig. 12). Complementation of *clv3-2* for immune marker gene expression in the SAM could be achieved with nanomolar range of exogenous CLV3p, reflecting the physiological relevance of the K_d for CLV3p and FLS2 interaction and CLV3p-FLS2 signalling in the SAM (Fig. 2e and Supplementary Fig. 10). The flg22 induction of these immune marker genes was observed in the SAM of *clv1-1*, *clv2-1* and *clv3-2* mutants, supporting the potency of MAMP signalling via FLS2 (Fig. 4a and Supplementary Fig. 11b). At 10–100 pM flg22, the immune marker gene induction levels were reduced in the SAM of *clv3-2* compared to those in the SAM of wild type and *clv1-1*, indicating the operation of both CLV3p-FLS2 and flg22-FLS2 signalling pathways in the SAM (Supplementary Fig. 13). The repression of *WUS* was triggered by CLV3p only in the SAM of wild type and *fls2-24*, but not *clv1-1* and *clv2-1* (Fig. 4b)^{1,2,4,17,19,23}.

Most notably, we have never detected the presence of a single live *P. syringae* pv. *tomato* DC3000–GFP bacterium in the SAM of wild-type seedlings in 11 infection experiments by visualizing bacteria proliferation using confocal microscopy for up to 4 days (Fig. 4c and Supplementary Fig. 14a). Because *P. syringae* pv. *tomato* DC3000–GFP could be easily visualized in the infected wild-type and *clv3-2* cotyledons (Supplementary Fig. 14b, d), the wild-type SAM appeared to exhibit differential immunity (Fig. 4c and Supplementary Fig. 14a). Notably, the SAM was no longer protected from *P. syringae* pv. *tomato* DC3000–GFP infection in *fls2-24* and *clv3-2* mutants (Fig. 4c and Supplementary Fig. 14c, e, f), supporting the important role of endogenous CLV3p in protection of the SAM through the FLS2-mediated innate immune signalling pathway (Fig. 4c and Supplementary Figs 1 and 14). To demonstrate further and quantify the proliferation and growth of *P. syringae* pv. *tomato* DC3000–GFP in the SAM of the *fls2-24* and *clv3-2* mutants observed after 2–4 days of infection, we counted the increasing bacteria numbers and carried out quantitative PCR analysis of the GFP DNA from the bacteria, both indicating the loss of the distinct SAM immunity (Fig. 4c and Supplementary Figs 14e and 15). There were higher numbers of *P. syringae* pv. *tomato* DC3000–GFP bacteria in the bigger *clv3-2* SAM with more cells of similar size but not bigger cells (Supplementary Fig. 14c, e). Importantly, *P. syringae* pv. *tomato* DC3000–GFP was found to be completely excluded from the similarly enlarged *clv1-1* and *clv2-1* SAM, in which CLV3p-FLS2 signalling remained active (Fig. 4c). On the basis of the qPCR analysis of GFP DNA derived only from *P. syringae* pv. *tomato* DC3000–GFP, nonspecific bacteria attachment background could be

estimated from the SAM tissue samples 1 h after bacteria–seedling co-cultivation. Consistent with confocal microscopic observations, active bacteria growth and proliferation were exclusively detected only in the SAM of the *clv3-2* and *fls2-24* mutants (Supplementary Fig. 15).

We have uncovered a surprising mechanism underlying the stem-cell-triggered immunity for pathogenic bacteria through CLV3p–FLS2 signalling (Supplementary Fig. 1). It will be interesting to examine the SAM protection from a variety of other pathogens. We propose that CLV3p is recognized by two distinct types of receptors involved in mostly non-overlapping functions in the SAM (Supplementary Fig. 1). The ‘constitutive’ immunity in the SAM resembles flg22 pre-treatment as a type of vaccination (before infection), which is more effective to confer protection against virulent pathogens such as *P. syringae* pv. *tomato* DC3000 possessing effectors to cripple MAMP signalling^{6,14}. It is surprising that CLV3p–FLS2 signalling seems to have evolved to provide constitutive immune protection in the SAM but avoid the penalty from potent growth suppression associated with MAMP signalling^{5,6,15,20}. A future challenge is to elucidate the precise differential downstream signalling events via the same receptor in response to different peptide ligands. Lacking the genes for antibodies and immune cell receptors in humans²⁹ to respond to new signals from diverse invaders, plant RLKs, displaying high polymorphism and fast evolution³⁰, may provide an alternative means to recognize self or non-self in a beneficial manner through constant selections in evolution. It will also be important to explore the roles of other secreted plant peptides and known or orphan receptors in innate immunity^{4,6,17,30}.

METHODS SUMMARY

Plasmid constructs. The constructs for expressing FLS2–haemagglutinin (HA) and BAK1–Flag for co-immunoprecipitation were reported previously¹⁴. LRR23b–HA and fls2-24–HA are mutant variants of the LRR domain of FLS2, which were generated by site-directed mutagenesis. The *fls2-24* allele is a Gly to Arg mutation in the 10th LRR domain of FLS2 and is insensitive to flg22 (ref. 25). LRR23b is mutated in two amino acids (Gln to Leu and Phe to Leu) in the 23rd LRR domain of FLS2 and lacks flg22 signalling²⁶.

Mesophyll protoplast transient assays. Protoplast isolation and transient expression assays were performed as previously described^{9,14}. For protein expression in co-immunoprecipitation and in-gel kinase assays, protoplasts were transfected with plasmid DNA and incubated for 6 h at room temperature. Then, peptides such as CLV3p and flg22 were added for 10 min to induce MAPK activation or FLS2–BAK1 interaction. For qRT–PCR analysis, protoplasts were treated with peptides for 1 h.

In-gel kinase assay and co-immunoprecipitation. Both experiments were performed according to procedures previously described^{9,14}. MBP (Invitrogen) was used as a substrate for endogenous MPK3 and MPK6 activation analysis⁹. For co-immunoprecipitation, FLS2–HA and BAK1–Flag were co-immunoprecipitated by an anti-Flag antibody (Sigma) and detected by an anti-HA antibody (Roche) in immunoblot analysis.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 March 2010; accepted 10 February 2011.

Published online 17 April 2011.

- Bäurle, I. & Laux, T. Apical meristems: the plant's fountain of youth. *Bioessays* **25**, 961–970 (2003).
- Scheres, B. Stem cells: a plant biology perspective. *Cell* **122**, 499–504 (2005).
- Hollings, M. Disease control through virus-free stock. *Annu. Rev. Phytopathol.* **3**, 367–396 (1965).
- Jun, J. H., Fiume, E. & Fletcher, J. C. The CLE family of plant polypeptide signaling molecules. *Cell. Mol. Life Sci.* **65**, 743–755 (2008).
- Gomez-Gomez, L. & Boller, T. FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Mol. Cell* **5**, 1003–1011 (2000).
- Boller, T. & Felix, G. A renaissance of elicitors: Perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* **60**, 379–406 (2009).

- Ishii, K. J., Koyama, S., Nakagawa, A., Coban, C. & Akira, S. Host innate immune receptors and beyond: Making sense of microbial infections. *Cell Host Microbe* **3**, 352–363 (2008).
- Curran, S. P., Wu, X., Riedel, C. G. & Ruvkun, G. A soma-to-germline transformation in long-lived *Caenorhabditis elegans* mutants. *Nature* **459**, 1079–1084 (2009).
- Asai, T. *et al.* MAP kinase signalling cascade in *Arabidopsis* innate immunity. *Nature* **415**, 977–983 (2002).
- Zipfel, C. *et al.* Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* **428**, 764–767 (2004).
- Yamamoto, C. *et al.* Rewiring mitogen-activated protein kinase cascade by positive feedback confers potato blight resistance. *Plant Physiol.* **140**, 681–692 (2006).
- Yadav, R. K., Girke, T., Pasala, S., Xie, M. & Reddy, G. V. Gene expression map of the *Arabidopsis* shoot apical meristem stem cell niche. *Proc. Natl Acad. Sci. USA* **106**, 4941–4946 (2009).
- Kondo, T. *et al.* A plant peptide encoded by CLV3 identified by in situ MALDI-TPF MS analysis. *Science* **313**, 845–848 (2006).
- Shan, L. *et al.* Bacterial effectors target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. *Cell Host Microbe* **4**, 17–27 (2008).
- Chinchilla, D. *et al.* A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* **448**, 497–500 (2007).
- Heese, A. *et al.* The receptor-like kinase SEERK3/BAK1 is a central regulator of innate immunity in plants. *Proc. Natl Acad. Sci. USA* **104**, 12217–12222 (2007).
- Butenko, M. A., Vie, A. K., Brembu, T., Aalen, R. B. & Bones, A. M. Plant peptides in signaling: looking for new partners. *Trends Plant Sci.* **14**, 255–263 (2009).
- Miwa, H. *et al.* The receptor-like kinase SOL2 mediates CLE signaling in *Arabidopsis*. *Plant Cell Physiol.* **49**, 1752–1757 (2008).
- Müller, R., Bleckmann, A. & Simon, R. The receptor kinase CORYNE of *Arabidopsis* transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *Plant Cell* **20**, 934–946 (2008).
- Zipfel, C. *et al.* Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell* **125**, 749–760 (2006).
- Fiers, M. *et al.* The CLAVATA3/ESR motif of CLAVATA3 is functionally independent from the nonconserved flanking sequences. *Plant Physiol.* **141**, 1284–1292 (2006).
- Felix, G., Duran, J. D., Volko, S. & Boller, T. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J.* **18**, 265–276 (1999).
- Bäurle, I. & Laux, T. Regulation of *WUSCHEL* transcription in the stem cell niche of the *Arabidopsis* shoot meristem. *Plant Cell* **17**, 2271–2280 (2005).
- Chinchilla, D., Bauer, Z., Regenass, M., Boller, T. & Felix, G. The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *Plant Cell* **18**, 465–476 (2006).
- Gómez-Gómez, L., Bauer, Z. & Boller, T. Both the extracellular leucine-rich repeat domain and the kinase activity of FLS2 are required for flagellin binding and signaling in *Arabidopsis*. *Plant Cell* **13**, 1155–1163 (2001).
- Dunning, F. M., Sun, W., Jansen, K. L., Helft, L. & Bent, A. F. Identification and mutational analysis of *Arabidopsis* FLS2 leucine-rich repeat domain residues that contribute to flagellin perception. *Plant Cell* **19**, 3297–3313 (2007).
- Bauer, Z., Gomez-Gomez, L., Boller, T. & Felix, G. Sensitivity of different ecotypes and mutants of *Arabidopsis thaliana* toward the bacterial elicitor flagellin correlates with the presence of receptor-binding sites. *J. Biol. Chem.* **276**, 45669–45676 (2001).
- Guo, Y., Han, L., Hymes, M., Denver, R. & Clark, S. E. CLAVATA2 forms a distinct CLE-binding receptor complex regulating *Arabidopsis* stem cell specification. *Plant J.* **63**, 889–900 (2010).
- Goldsby, R. A., Kindt, T. J. & Osborne, B. A. *Kuby Immunology* (W. H. Freeman, New York, 2006).
- Clark, R. M. *et al.* Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* **317**, 338–342 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Laux for the *pWUS::GUS* line, Y. Matsubayashi for the purified Ara-CLV3 peptide and advice, A. Collmer for the GFP-labelled *P. syringae* pv. *tomato* DC3000, L. Shan and P. He for constructs, G. Tena, F. Ausubel, J. Bush, J. Plotnikova and ABRC for mutant seeds and bacterial strains, and Y. Xiong, M. McCormack, Y. Niu, M. Ramon and J. Li for critically reading of the manuscript. Funding was provided by the NSF, NIH and the MGH CCIB fund to J.S.

Author Contributions H.L. and J.S. initiated the project and designed the experiments; H.L. carried out the experiments and prepared the data with assistance from O.-K.C.; H.L. and J.S. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.L. (hrlee@molbio.mgh.harvard.edu) or J.S. (sheen@molbio.mgh.harvard.edu).

METHODS

Plasmid constructs. The constructs for expressing FLS2-haemagglutinin (HA) and BAK1-Flag for co-immunoprecipitation were reported previously¹⁴. LRR23b-HA and fls2-24-HA are mutant variants of the LRR domain of FLS2, which were generated by site-directed mutagenesis. The *fls2-24* allele is a Gly to Arg mutation in the 10th LRR domain of FLS2 and is insensitive to flg22 (ref. 25). LRR23b is mutated in two amino acids (Gln to Leu and Phe to Leu) in the 23rd LRR domain of FLS2 and lacks flg22 signalling²⁶.

Mesophyll protoplast transient assays. Protoplast isolation and transient expression assays were performed as previously described^{9,14}. For protein expression in co-immunoprecipitation and in-gel kinase assays, protoplasts were transfected with plasmid DNA and incubated for 6 h at room temperature. Then, peptides such as CLV3p and flg22 were added for 10 min to induce MAPK activation or FLS2-BAK1 interaction. For qRT-PCR analysis, protoplasts were treated with peptides for 1 h.

In-gel kinase assay and co-immunoprecipitation. Both experiments were performed according to procedures previously described^{9,14}. MBP (Invitrogen) was used as a substrate for endogenous MPK3 and MPK6 activation analysis⁹. For co-immunoprecipitation, FLS2-HA and BAK1-Flag were co-immunoprecipitated by an anti-Flag antibody (Sigma) and detected by an anti-HA antibody (Roche) in immunoblot analysis.

Plant materials and growth conditions. Col-0 and Ler were used as wild-type *Arabidopsis* plants in this study. The *bak1-4*, *efr-1*, and *fls2* (Salk_141277) mutants are in the Col-0 background^{14,15,20} and *clv1-1*, *clv2-1*, *clv3-2* and *fls2-24* are in Ler background^{5,31,32}. The *clv1-1* mutation is in the kinase domain and represents a dominant-negative allele³³. The mutation of *clv2-1* causes the early stop codon at the 33rd residue and is a null allele³⁴. The γ -ray-induced *clv3-2* is a presumed null allele^{33,35}. Wild-type and mutant plants were grown on soil at 23 °C, 65% humidity, and 75 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity under 12 h light/12 h dark photoperiod conditions for 4 weeks before mesophyll protoplast isolation³⁶. For liquid culture of *Arabidopsis* seedlings, seeds were germinated and grew in 6-well plates containing 1 ml of liquid medium (0.5 \times MS and 0.5% sucrose, pH 5.8 adjusted with KOH). For GUS assay in the SAM and qRT-PCR analysis using SAM tissues, seedlings were grown for 8 days. For growth inhibition analysis by flg22 and CLV3p, seedlings were grown at the same condition for 8 days (Fig. 3a, b and Supplementary Fig. 8b) or 6 days (Fig. 3c–e and Supplementary Fig. 7b).

qRT-PCR analysis. Total RNA was isolated from protoplasts or SAM tissues with TRIzol reagent (Invitrogen). To harvest SAM tissues, seedlings were instantaneously frozen by liquid nitrogen in the mortar pre-chilled on the dry ice. Cotyledons, hypocotyls and roots were removed using fine forceps. The purity of the harvested SAM tissue was confirmed by SAM-specific marker genes and the absence of marker genes not expressed in the SAM (Supplementary Fig. 2)¹². First-strand cDNA was synthesized from 1 μg of total RNA with M-MLV reverse transcriptase (Promega). All qRT-PCR analyses were performed by CFX96 real-time PCR detection system with iQ SYBR green supermix (Bio-Rad). *ACT2* (*ACTIN2*, At3g18780) was used as a control gene.

GUS staining. GUS staining was performed as described²³. Plants were fixed in 90% cold acetone for 20 min and rinsed twice in staining buffer without X-Gluc (5-bromo-4-chloro-3-indoxyl-beta-D-glucuronide). Samples were infiltrated with staining buffer (100 mM NaPO₄ buffer, pH 7.0; 0.5 mM ferrocyanide; 0.5 mM ferricyanide; 0.1% Triton X-100; 10 mM EDTA, pH 8.0; 1 mM X-Gluc (Gold Biotechnology)) under vacuum for 10 min and incubated at 37 °C for 3 h. After then, staining buffer was removed and dehydrated up to 70% ethanol. Microscopic analysis was carried out with Leica DFC 500 camera mounted on Leica MZ16F.

Bacterial infection assay in the SAM. Nine seeds were sowed in 1-ml liquid medium in 6-well plates and grown under constant light (50–65 $\mu\text{mol m}^{-2} \text{s}^{-1}$) at 25–27 °C without shaking. *P. syringae* pv. *tomato* DC3000-GFP culture was grown in KB liquid medium (50 $\mu\text{g ml}^{-1}$ of rifampicin and 15 $\mu\text{g ml}^{-1}$ of tetracycline) with shaking at 28 °C. Overnight cultured *P. syringae* pv. *tomato* DC3000-GFP were washed twice with water and diluted to an optical density at 600 nm (OD_{600}) of 0.02. Diluted *P. syringae* pv. *tomato* DC3000-GFP (50 μl) was added into the liquid medium with 2-day-old seedlings. Plants and bacteria were co-cultivated with gentle shaking (50 r.p.m.) for 2, 3 or 4 days under constant light. For observation of bacteria-infected SAM, co-cultivated seedlings were washed with 70% ethanol twice and rinsed twice with water. All seedlings were placed on glass slides with 1 ml of water, and squashed gently by coverslips for bacterial observation using a confocal laser-scanning microscope (Leica TCS-NT). The experiment was repeated at least five times with similar results in Ler, *fls2-24*, *clv3-2*, *clv1-1* and *clv2-1* (Fig. 4c). For the qPCR analysis of DC3000-GFP, diluted *P. syringae* pv. *tomato* DC3000-GFP ($\text{OD}_{600} = 0.5$, 200 μl) was co-cultivated with wild-type or mutant seedlings for 3 or 4 days. After washing and rinsing seedlings twice, tissues from five SAMs were harvested and ground in 100 μl of water.

Control experiments were conducted at 0 days after co-cultivation to determine nonspecific bacterial attachment 1 h after co-cultivation (black bar). The experiment was repeated three times with similar results. The GFP level determined by qPCR was correlated with specific bacterial growth, and normalized based on the *Arabidopsis* *ACT2* gene in the SAM tissues.

Seedling pathogen assay. The protocol was modified from that previously described^{37,38}. Nine seeds were sowed in 6-well plates containing 1 ml of liquid medium. Plants were grown under constant light at 25–27 °C without shaking. *P. syringae* pv. *tomato* DC3000 culture was grown in KB liquid medium (50 $\mu\text{g ml}^{-1}$ of rifampicin) with shaking at 28 °C. Overnight cultured *P. syringae* pv. *tomato* DC3000 was washed twice with water and diluted to $\text{OD}_{600} = 0.02$, 1×10^7 c.f.u. ml^{-1} . After 6 days of seedling growth, 50 μl of diluted *P. syringae* pv. *tomato* DC3000 was added into 1 ml of fresh 0.5 \times MS liquid medium without sucrose. After adding *P. syringae* pv. *tomato* DC3000, plates containing plants and bacteria were co-cultivated with gentle shaking (<50 r.p.m.) for 1 day under constant light (50–65 $\mu\text{mol m}^{-2} \text{s}^{-1}$). For CLV3p- and flg22-induced immunity, seedlings were treated with 1 μM or 10 μM of CLV3p or 100 nM of flg22 peptide 1 day before adding *P. syringae* pv. *tomato* DC3000. For bacterial counting, co-cultivated seedlings were washed with 70% ethanol twice and rinsed twice with water. Then, three seedlings were put into each of three 1.5-ml tubes containing 100 μl of water and ground by a hand drill and blue pestles. After preparing serial dilutions, 10 μl of diluted bacteria (from 10^{-3} to 10^{-5}) was spread on the KB plates and incubated for 2 days at 28 °C before counting.

Assay for CLV3p-mediated SAM arrest. Twenty seeds of Col-0 and *fls2* mutant were sowed in Petri dishes (100 \times 25 mm) containing 10 ml of liquid medium with or without 1 μM CLV3p. Plants were grown in a growth chamber at 23 °C under short day condition (8 h light/16 h dark, 75 $\mu\text{mol m}^{-2} \text{s}^{-1}$) for 4 weeks. The SAM tissues were cut and fixed using 4% (w/v) paraformaldehyde/4% (v/v) DMSO at 4 °C overnight. Collected samples were dehydrated through ethanol series (30%, 50%, 70%, 95% for 1 h in each step) at 4 °C and stained with 0.1% Eosin Y (Sigma) in 100% ethanol at 4 °C overnight. Ethanol was changed through histoclear series (50% ethanol: 50% histoclear, 100% histoclear, 100% histoclear for 1 h in each step) (National Diagnostics). Histoclear was then gradually changed with melted paraffin (Fisher Scientific) in a 60 °C chamber. Replacement of freshly melted paraffin was performed for 4 days. Paraffin-embedded tissues were poured into the mould and adjusted in appropriated position. Section was carried out with a rotary microtome (Leica RM2255) at 8- μm thickness. Sectioned ribbons were placed on poly-prep glass slides (Sigma) with pre-warm water and incubated on a slide warmer (Fisher Scientific) at 42 °C overnight. For meristem staining, paraffin of sections was removed by 100% histoclear twice for 10 min. Sectioned tissues were hydrated through reverse ethanol series (100%, 70%, 30% ethanol and water for 2 min in each step). Sections were stained with 0.1% Giemsa (Sigma) for 5 min and rinsed briefly with water. Stained sections were dehydrated through ethanol series (2 min in each step) and transferred to 100% histoclear for 2 min. For microscopic analysis, samples were dried and mounted with Cytoseal 60 (Richard-Allan Scientific) before taking pictures using Leica DM5000B.

Direct binding assay. The protocol was modified from that previously described^{27,28,39,40}. For preparation of receptor proteins, FLS2-HA, fls2-24-HA, LRR23b-HA and BR11-HA were expressed in *fls2* protoplasts (2.5×10^5 cells) for 6 h. Protoplasts were harvested and then re-suspended in 500 μl of binding buffer (25 mM MES/KOH, pH 5.8, 3 mM MgCl₂, 10 mM NaCl) with 2 mM DTT and protease inhibitor (Roche). Cells were vigorously mixed by vortex and kept on ice for 5 min, and centrifuged at 10,000g for 20 min to yield pellet. The pellets were re-suspended in 100 μl of binding buffer and the protein concentration from re-suspended extract was measured by a NanoDrop 1000 spectrophotometer. This extract contained 50 μg of total protein. For the binding assay, re-suspended extracts (100 μl) in binding buffer were mixed with ¹²⁵I-Tyr-CLV3p (100 fmol in each sample) without or with unlabelled Tyr-CLV3p as a competitor for 30 min on ice. The average of specific radioactivity in five different batches of ¹²⁵I-Tyr-CLV3p was 2,023.52 Ci mmol⁻¹. ¹²⁵I-Tyr-CLV3p bound extracts were collected by a vacuum filtration system through glass fibre filters (Macherey-Nagel MN GF-2, 2.5-cm diameter), which were pre-incubated with 1% BSA, 1% bacto-tryptone, 1% bacto-pepton in binding buffer. Filters were washed with 15 ml of cold binding buffer and the retained radioactivity was determined by a gamma counter Beckman LS6500. Specific binding was measured by subtracting nonspecific binding (with 10–20 μM unlabelled Tyr-CLV3p competitor) from total binding (without competitor). In the saturation binding assay, the nonspecific binding in the presence of 20 μM unlabelled Tyr-CLV3p competitor showed linear increase with 1 to 100 nM ¹²⁵I-Tyr-CLV3p and accounted for 40–60% of total binding, which was similar to the range of nonspecific binding observed with ¹²⁵I-Tyr-flg22 binding to intact cells (Fig. 3a, Bauer *et al.*, 2001)²⁷. The dissociation constant (K_d), Scatchard plot and the nonlinear regression analysis of competition assay were presented using the Prism 5 program (GraphPad). From the saturation data

of $B_{\max} = 0.53 \text{ pmol mg}^{-1} \text{ protein}$ and $50 \mu\text{g protein}$ from 2.5×10^5 cells, 6.36×10^4 ^{125}I -Tyr-CLV3p binding sites per cell could be estimated in total cell extracts. Specific CLV3p binding to FLS2 was similar to CLV3p binding to CLV1 and CLV2, the established CLV3p receptors^{28,39}.

31. Koornneef, M. *et al.* Linkage map of *Arabidopsis thaliana*. *J. Hered.* **74**, 265–272 (1983).
32. Clark, S. E., Running, M. P. & Meyerowitz, E. M. CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. *Development* **121**, 2057–2067 (1995).
33. Diévar, A. *et al.* CLAVATA1 dominant-negative alleles reveal functional overlap between multiple receptor kinases that regulate meristem and organ development. *Plant Cell* **15**, 1198–1211 (2003).
34. Jeong, S., Trotochaud, A. E. & Clark, S. E. The *Arabidopsis* CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* **11**, 1925–1933 (1999).
35. Fletcher, J. C., Brand, U., Running, M. P., Simon, R. & Meyerowitz, E. M. Signaling of cell fate decisions by CLAVATA3 in *Arabidopsis* shoot meristems. *Science* **283**, 1911–1914 (1999).
36. Yoo, S. D., Cho, Y. H. & Sheen, J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols* **2**, 1565–1572 (2007).
37. Schreiber, K., Ckurshumova, W., Peek, J. & Desveaux, D. A high-throughput chemical screen for resistance to *Pseudomonas syringae* in *Arabidopsis*. *Plant J.* **54**, 522–531 (2008).
38. Boudsocq, M. *et al.* Differential innate immune signaling via Ca^{2+} sensor protein kinases. *Nature* **464**, 418–422 (2010).
39. Ogawa, M., Shinohara, H., Sakagami, Y. & Matsubayashi, Y. *Arabidopsis* CLV3 peptide directly binds CLV1 ectodomain. *Science* **319**, 294 (2008).
40. Ohya, K., Shinohara, H., Ogawa-Ohnishi, M. & Matsubayashi, Y. A glycopeptide regulating stem cell fate in *Arabidopsis thaliana*. *Nature Chem. Biol.* **5**, 578–580 (2009).

A novel protein family mediates Casparian strip formation in the endodermis

Daniele Roppolo¹, Bert De Rybel^{2†}, Valérie Dénervaud Tendon¹, Alexandre Pfister¹, Julien Alassimone¹, Joop E. M. Vermeer¹, Misako Yamazaki¹, York-Dieter Stierhof³, Tom Beeckman² & Niko Geldner¹

Polarized epithelia are fundamental to multicellular life. In animal epithelia, conserved junctional complexes establish membrane diffusion barriers, cellular adherence and sealing of the extracellular space¹. Plant cellular barriers are of independent evolutionary origin. The root endodermis strongly resembles a polarized epithelium and functions in nutrient uptake and stress resistance². Its defining features are the Casparian strips, belts of specialized cell wall material that generate an extracellular diffusion barrier². The mechanisms localizing Casparian strips are unknown. Here we identify and characterize a family of transmembrane proteins of previously unknown function. These ‘CASPs’ (Casparian strip membrane domain proteins) specifically mark a membrane domain that predicts the formation of Casparian strips. CASP1 displays numerous features required for a constituent of a plant junctional complex: it forms complexes with other CASPs; it becomes immobile upon localization; and it sediments like a large polymer. CASP double mutants display disorganized Casparian strips, demonstrating a role for CASPs in structuring and localizing this cell wall modification. To our knowledge, CASPs are the first molecular factors that are shown to establish a plasma membrane and extracellular diffusion barrier in plants, and represent a novel way of epithelial barrier formation in eukaryotes.

In 1865 Robert Caspary described a novel cell type surrounding the vascular cylinder of roots³. Caspary named this layer ‘Schuttscheide’, protective sheath, but during the twentieth century it was termed endodermis, the ‘inner skin’ of plants. Its defining, belt-like cell wall thickenings became known as Casparian strips. The ligno-suberic polymers of the Casparian strips impregnate the primary cell walls between cells, providing a localized diffusion barrier^{4,5}. The appearance of Casparian strips in evolution coincides with that of vascular tissues, and a huge body of physiological work proposes numerous functions for Casparian strips^{2,6–8}. Despite their importance, and our significant insights into endodermal specification^{9–11}, we remain ignorant about the molecules that govern endodermal differentiation. Recently, polarized distribution of transporters within the endodermis was demonstrated^{12–14}. Moreover, a central domain, underlying the Casparian strip (Casparian strip membrane domain, CSD) was shown to separate the polar domains and prevent lateral diffusion¹³. This CSD displays tight matrix attachment and appears very electron dense in electron micrographs¹⁵. Despite obvious functional resemblances between endodermis and animal epithelia¹⁶, plants lack most components of animal junctional complexes, and plant cell wall thickness does not allow for direct, protein-mediated cell–cell contact, crucial for animal tight junctions. Therefore, sealing must rely on the coordinated deposition of hydrophobic cell wall material by some completely unknown mechanism.

In *Arabidopsis thaliana*, Casparian strips are endodermis-specific structures. We therefore searched microarrays¹⁷ for endodermis-enriched genes, the products of which were predicted to be secreted

or plasma-membrane localized. Five proteins (>45% amino acid identity) were identified that belong to the *Arabidopsis* ‘uncharacterized protein family’ UPF0497 (38 members, Supplementary Fig. 1a). Family members have a predicted topology of four-membrane spans with cytosolic amino and carboxy termini and conserved extracellular loops (Fig. 1a and Supplementary Fig. 1b). Reporter lines showed strong and specific transcription in the endodermis, starting in the elongation zone (Supplementary Fig. 2a). Fluorescent protein fusions of all five proteins under endogenous or endodermis-expressing *SCARECROW* (*SCR*) promoter¹⁸ showed very restricted localization within the plasma membrane, coinciding with the position of the CSD (Fig. 1b–d and Supplementary Fig. 2b, c). The CASP1–GFP signals are strictly complementary with those of markers for peripheral (outer) and central (inner) plasma membrane domains¹³ (Fig. 1e, f and Supplementary Fig. 3). As the first proteins marking the CSD, we termed them Casparian strip membrane domain proteins 1 to 5 (CASP1–5). In differentiated endodermal cells, Casparian strips form a supracellular network of crosslinked cell walls¹⁹. Three-dimensional reconstructions of CASP1–GFP revealed a markedly similar network of precisely aligned belts of CASP1–GFP signals forming a cylindrical network around the stele (Supplementary Movie 1). By immuno-electron microscopy, we found that width and position of the CASP1–GFP signals at the plasma membrane precisely coincided with the Casparian strip itself (Fig. 1g, h and Supplementary Fig. 4). CASP1–GFP signal was aligned between neighbouring cells, but not in direct contact. In addition, CASP1 signals at the membrane also matched the zone of adherence to the Casparian strip.

CASPs could simply be associated proteins of an established Casparian strip network. Alternatively, CASP localization could precede and determine the site of Casparian strip deposition. To discriminate between these possibilities, we developmentally staged the onset of CASP1 expression, with respect to CSD appearance and Casparian strip diffusion barrier formation (Supplementary Fig. 5). In *Arabidopsis*, strict division patterns allow staging by counting the cellular distance from the meristem¹³. We found that *CASP1* transcription precedes formation of the CSD by 2.6 ± 0.8 cells, visualized as a zone of protein exclusion. This matches the difference between initial CASP1 protein accumulation and its final localization (2.0 cells). Moreover, CASP1 localization precedes the establishment of a functional diffusion barrier by 4.8 ± 1.5 cells. Thus, CASPs precede Casparian strip establishment and are early markers of CSD formation. Moreover, live imaging of early CASP1–GFP accumulation indicates that CASPs could be involved in generating the CSD (Fig. 2a and Supplementary Movie 2). CASP1–GFP shows an initially uniform distribution at the plasma membrane. Signal strength then increases and accumulation at sites of incipient CSDs is observed. Localization in other plasma membrane regions gradually disappears and signals at the forming CSDs become stronger and more defined. Figure 2a shows a representative time lapse of this progression, taking place within 2 h (Supplementary Movie 2). Intriguingly, imaging

¹Department of Plant Molecular Biology, Biophore, UNIL-Sorge, University of Lausanne, 1015 Lausanne, Switzerland. ²Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Genetics, Ghent University, Technologiepark 927, 9052 Ghent, Belgium. ³ZMBP, Mikroskopie, Universität Tübingen, Auf der Morgenstelle 5, 72076 Tübingen, Germany. [†]Present address: Wageningen University, Laboratory of Biochemistry, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands.

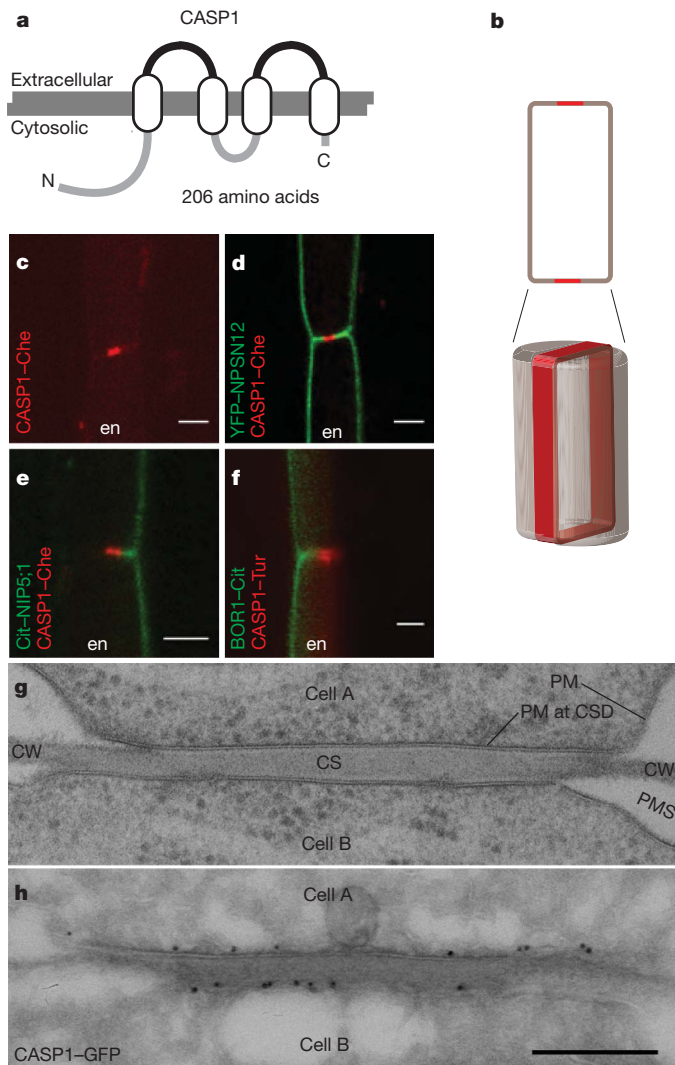


Figure 1 | A new protein family localizes to the Casparian strip membrane domain. **a**, Predicted topology of CASP1. **b**, Cartoon of the observed dots in an optical section that represent a longitudinal band in three dimensions. **c–f**, CASP1 (red) is confined to the CSD (**c**), is complementary to NPSN12 (**d**), and outer (NIP5;1, **e**) and inner (BOR1(Y373A/Y398A/Y405A), **f**) markers. Che, mCherry; Cit, citrine; en, endodermis; Tur, mTurquoise. **g**, Casparian strips (CS) appear homogeneous in electron micrographs. CSD adheres to Casparian strips. CW, cell wall; PM, plasma membrane; PMS, space generated by plasmolysis. **h**, Immunogold-electron micrograph of CASP1–GFP. Gold particles reside at either side of the Casparian strip. Scale bars: **c–f**, 5 μ m; **h**, **i**, 250 nm.

of the cell surface reveals that localization proceeds through a ‘string-of-pearls’ stage where CASP1–GFP patches appear along the equatorial line of the cell, gradually coalescing into a continuous band (Fig. 2b and Supplementary Movie 3). Initial random distribution, gradual relocation and patching at the future site of Casparian-strip formation all indicate that CASPs are not locating to a previously established domain but are associated with its formation. When expressed ectopically, most CASPs reach the plasma membrane, but none accumulates in CSD-like structures, indicating that unknown endodermis-specific factors, possibly other CASPs, are necessary for their localization (Supplementary Fig. 6). Re-localization of CASP1–GFP should be associated with endocytosis and secretion of the protein. Indeed, CASP1–GFP accumulates in endosomal aggregates in response to brefeldin A (BFA), indicative of active endocytosis (Fig. 2c)²⁰. However, this accumulation eventually disappears in differentiated cells where CASP1 localization has become confined to the CSD (Fig. 2d). Other proteins continued to

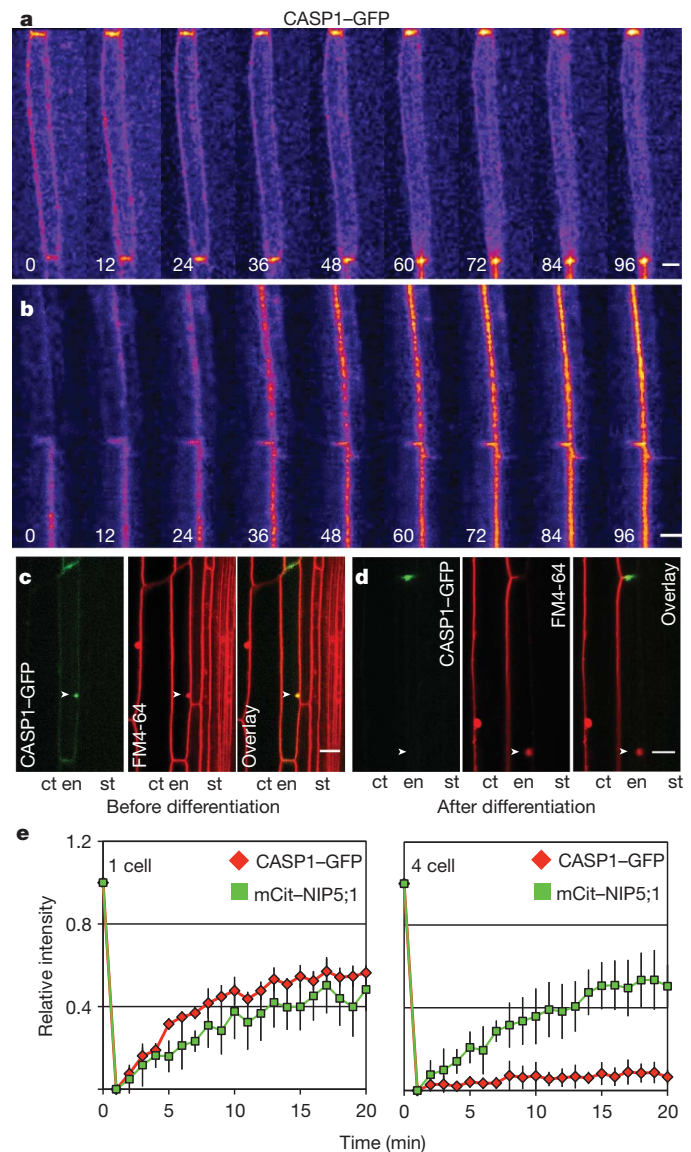


Figure 2 | CASP1 protein gradually localizes to a stable central domain. **a, b**, CASP1–GFP time-lapse images. Numbers at the bottom of panels **a** and **b** indicate time in minutes. Central (**a**) and surface (**b**) optical cuts are shown. **c, d**, Effects of BFA on CASP1 localization. CASP1–GFP endosomal aggregates form in early (**c**, yellow in overlay) but not in late stages (**d**, red in overlay). Accumulation of FM4-64 (red) indicates ongoing BFA-sensitive recycling. Arrowheads point to endosomal aggregates. ct, cortex; en, endodermis; st, stela. Scale bars: 10 μ m. **e**, FRAP experiments with CASP1–GFP (red). Recovery is observed for CASP1–GFP (red) in the first expressing cell (1 cell), but not at later stages (4 cell). mCitrine–NIP5;1 (green) is a control protein. $n = 7–9$ independent assays for each series, error bars = s.d.

display BFA sensitivity (Supplementary Fig. 7). This suggests very low endocytosis of localized CASP1–GFP. Furthermore, CASP1–GFP only shows measurable rates of fluorescence recovery after photobleaching (FRAP) in early-expressing cells, but becomes immobile upon accumulation at the position of the CSD, strongly contrasting with other endodermal plasma membrane markers (Fig. 2e). This indicates that localized CASP1 has very low rates of lateral diffusion. Together, this suggests an extensive scaffolding and/or matrix interaction of CASP1. This behaviour of CASP1 perfectly matches the characteristics of the CSD itself, which we have shown to act as a molecular fence, excluding other membrane proteins and blocking diffusion between central and peripheral plasma membrane domains¹³.

A straightforward explanation for the immobility of CASP1–GFP and the presence of gradually fusing patches would be that CASPs

form polymeric lattices within the plane of the membrane. When ectopically expressed, GFP fusions of CASP1, CASP2, CASP3 and CASP4 accumulate at the plasma membrane in a non-localized fashion, occasionally labelling intracellular compartments (Fig. 3a and Supplementary Fig. 6). CASP5-GFP, however, shows strong accumulation in large, irregular, intracellular compartments (Fig. 3b). By structural and immunogold electron microscopy we found that misexpressed CASP5-GFP induces altered domains of endoplasmic reticulum (ER) in which it accumulates (Fig. 3c–e and Supplementary Fig. 8). Extensive, flattened ER cisternae often organized into large, multilamellar structures. This CASP5-induced re-organization and aggregation of membranes could best be explained by CASP5 forming extensive protein scaffolds within the ER. Clearly, our findings indicate that CASP5, especially, requires endodermis-specific factors, possibly other CASPs, for its exit from the ER. Using CASP1-GFP immunoprecipitation/mass spectrometry (IP/MS) analysis, we then tested whether

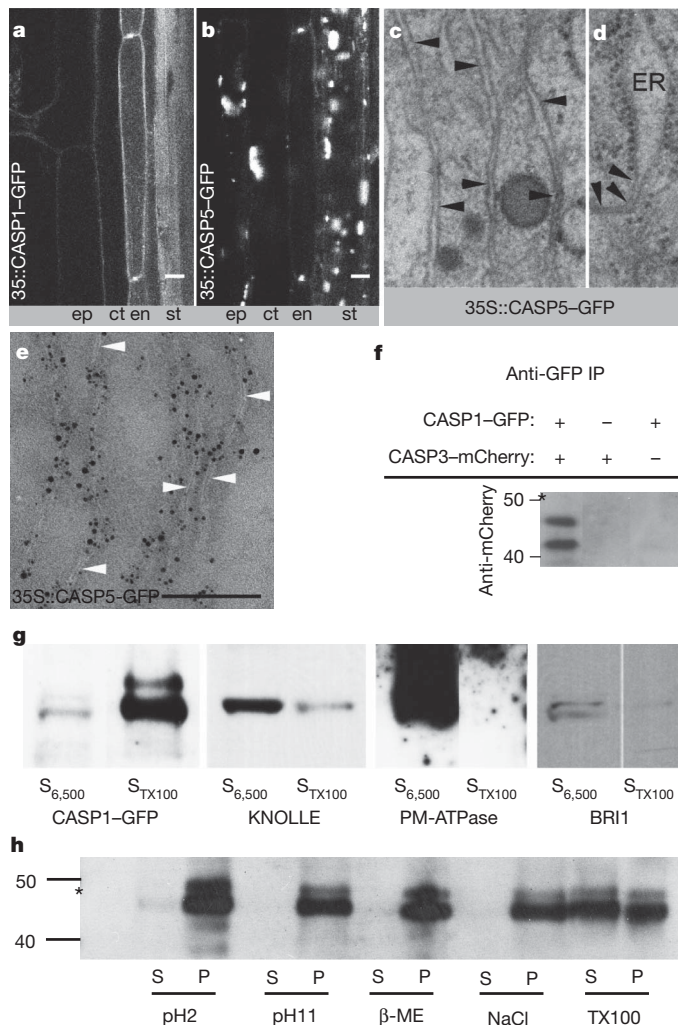


Figure 3 | CASP proteins interact with each other and show behaviour of large, polymeric protein assemblies. **a, b**, Ectopically expressed CASP1-GFP (**a**) and CASP5-GFP (**b**). ct, cortex; en, endodermis; ep, epidermis; st, stele. Scale bars: 10 μm. **c–e**, Epon embedded (**c, d**) and immunogold labelled (**e**) 35S::CASP5-GFP. Aberrant ER cisternae (black arrowheads, **c, d**) are labelled with anti-GFP (**e**, white arrowheads showing parallel bilayers). Scale bar: 250 nm. **f**, CASP1-GFP/CASP3-mCherry interaction. CASP3-mCherry is observed as lower molecular mass forms (asterisk indicates expected molecular mass of 52 kDa). **g**, CASP1-GFP in low-speed pellets. S_{6,500}, supernatant; S_{TX100}, supernatant after solubilization with Triton X-100. KNOLLE, PM-ATPase and BRI1 show contrasting fractionation. **h**, Solubilization with Na-citrate buffer (pH2), Na₂CO₃ buffer (pH11), 5% β-mercaptoethanol (β-ME), 1 M NaCl and 5% Triton X-100. P, pellet; S, supernatant.

CASPs reside in complexes with each other. By immunoprecipitations with anti-GFP antibodies, using a CASP1::CASP1-GFP line, we identified CASP3 as an interacting protein (Supplementary Fig. 9) and confirmed this by co-immunoprecipitation using a CASP1::CASP1-GFP, SCR::CASP3-mCherry double marker line (Fig. 3f). Similarly, we demonstrated interactions of multiple other CASP pairs (Supplementary Fig. 10), supporting the idea of extensive multivalent interactions between CASPs. These interactions may contribute to CASP localization at the CSD, by trapping newly arriving CASPs through polymerization.

A polymeric lattice of CASPs that surrounds an entire cell should have particular physical properties, because of its considerable size and mass. We therefore compared solubility of CASP1-GFP with that of other plasma membrane proteins and found that under native extraction conditions, the large majority of CASP1-GFP was not recovered in the microsomal fraction, as is the case for most transmembrane proteins (Fig. 3g). Instead, CASP1-GFP was found associated with low-speed pellets from which it could only be eluted by strong detergents (Fig. 3h and Supplementary Fig. 11). Polymer formation could easily account for this unusual behaviour. The conditions needed for CASP1-GFP release from pellets are markedly similar to those of Casparian-strip-associated proteins whose purification was attempted previously²¹. Clearly, fractionation into low-speed pellets could also be due to association of CASPs with the Casparian strip, and polymerization and cell wall association are obviously not mutually exclusive.

Our data indicate that CASPs establish the CSD, possibly guided by some unidentified earlier positional cue. This, in turn, determines the localization of the Casparian strip itself and the establishment of an apoplastic barrier¹³. To demonstrate this, we analysed available mutants for the small CASP genes. Neither insertion lines nor targeting induced local lesions in genomes (TILLING) stop mutations were found for CASP2 and CASP4 (Supplementary Figs 1c and 12). The best *casp1* allele identified was an insertion line that shows 85% knockdown of messenger RNA levels (*casp1-1*). Knockout alleles were identified for CASP3 (*casp3-1* and *casp3-2*) and a strong knockdown for CASP5 (*casp5-1*). Casparian strips are easily visualized by their intrinsic

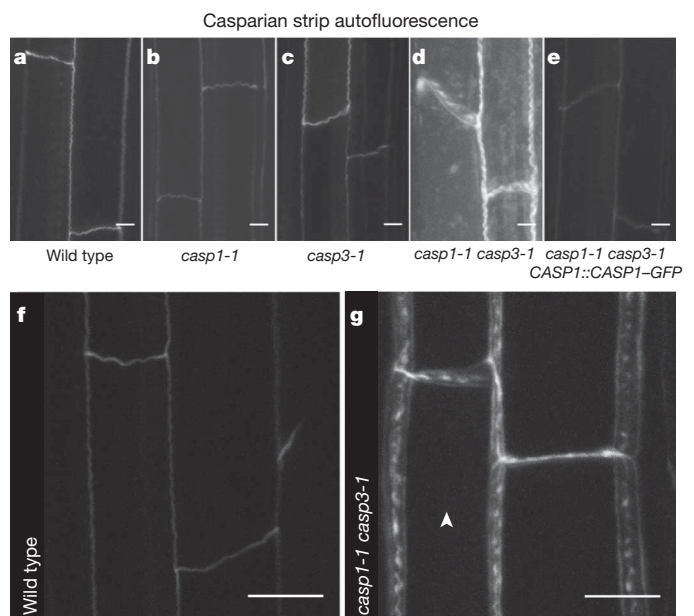


Figure 4 | CASP proteins are necessary for the correct structure and localization of the Casparian strips. **a–e**, Surface view of Casparian strip network, detected as autofluorescence after clearing with two-photon excitation. Z-stack projections (0.38 μm z-resolution), taken with identical settings. **f, g**, Higher magnification images of wild type (**f**) and *casp1-1 casp3-1* (**g**) with settings optimized for image quality. Scale bars: 10 μm.

autofluorescence after clearing of the roots¹³. Using this test, we did not observe any difference in Casparian strip formation between single *casp1* or *casp3* mutants and wild type (Fig. 4a–c). Absence of a phenotype in single mutants could be expected considering the similarity and identical expression and localization pattern of the five family members. RNA interference approaches were efficient in reducing overall mRNA levels of CASPs, but not the protein levels of CASP–GFP fusions in differentiating cells (Supplementary Fig. 12). Yet, combining *casp1-1* and *casp3-1* insertion mutants did yield obvious defects in Casparian strip formation. Instead of the strictly confined signal of wild-type Casparian strips, we noticed a much stronger and more diffuse autofluorescence in the double mutant (Fig. 4d), which was complemented by a CASP1::CASP1–GFP transgene (Fig. 4e). Three-dimensional reconstructions revealed that *casp1-1 casp3-1* mutants deposit autofluorescent cell wall material everywhere in the transversal and anticlinal walls with some preference for the cell corners. At the position of the Casparian strip, irregular, non-contiguous patches of signal are observed (Fig. 4f, g and Supplementary Movie 4). This demonstrates that CASPs are needed for the deposition of Casparian strip material into a contiguous, centrally positioned ring, but are not required for its synthesis or polymerization. This aberrantly structured Casparian strip apparently remains functional, as the apoplastic tracer propidium iodide is still blocked at the endodermis in the double mutant. Consistently, no obvious growth defects are observed in *casp1-1 casp3-1* plants. We expect to see strong phenotypes only in quintuple mutants that eliminate the activity of all CASP family members.

Our analysis indicates that CASPs are central, probably structural, components of the Casparian strip membrane domain, forming a tight, polymeric scaffold within the membrane. We speculate that the CASPs serve as a platform to localize and immobilize cell wall biosynthetic enzymes and that matrix adhesion is mediated either through direct interaction or interacting proteins. Identification of CASP-regulatory factors and associated proteins will allow a mechanistic understanding of Casparian strip formation, and enable us to revisit long-standing concepts about root nutrient uptake through specific manipulation of Casparian strip formation. CASP-related proteins probably perform related functions in other cell types. Their common role might be to form membrane protein platforms for the localization of cell-wall-modifying activities or localized adhesion. Mechanisms of localized wall deposition and matrix adhesion in plants remain poorly understood processes, and the identification and characterization of the CASPs and their related proteins now provide an intriguing new avenue for their mechanistic dissection.

METHODS SUMMARY

For FRAP, 1-h treatment with 50 mM 2-deoxyglucose/0.02% sodium azide and bleaching on 2 μm^2 was carried out. For time lapse, seedlings were transferred to Lab-Tek II chambered coverglass and covered with agar block. MBF ImageJ bundle was used for analysis. Native protein extraction was done with 100 mg of roots, 200 μl extraction buffer ((in mM): 50 HEPES, pH 7.9, 300 sucrose, 150 NaCl, 5 EDTA, 10 CH_3COOK , 2.5 \times Roche Complete, 1 PMSF); centrifugation for 5 min at 6,500g, 4 °C; washes, 3 \times in extraction buffer. Solubilization was performed with extraction buffer plus detergents. For immunoprecipitation, pellet obtained after centrifugation at 6,500g ($P_{6,500}$) from 450 mg of roots was treated for 30 min on ice with 450 μl extraction buffer plus 1% CHAPS, and then spun for 30 min at 19,000g at 4 °C. Then, 410 μl were incubated on ice for 30 min with 50 μl of μMACS anti-GFP microbeads. Beads were washed on column with 4 \times 200 μl , 1 \times 100 μl extraction buffer + 1% CHAPS. Elution was performed at 95 °C with 50 μl 1 \times μMACS extraction buffer. For IP-MS, 2 g of roots were used; $P_{6,500}$ was re-suspended in extraction buffer + 0.5% Triton X-100. Anti-GFP (1:1,000) was from Invitrogen, anti-DsRed (1:1,000) was from Clontech. Sequences have been deposited in GenBank (HQ699533–50). Inhibitor and tracer treatments¹³, high-pressure freezing²², immunogold localization²³ and IP-MS²⁴ were carried out as described previously.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 December 2010; accepted 24 March 2011.

- Chiba, H., Osanai, M., Murata, M., Kojima, T. & Sawada, N. Transmembrane proteins of tight junctions. *Biochim. Biophys. Acta* **1778**, 588–600 (2008).
- Enstone, D. E., Peterson, C. A. & Ma, F. S. Root endodermis and exodermis: Structure, function, and responses to the environment. *J. Plant Growth Regul.* **21**, 335–351 (2002).
- Caspary, R. Bemerkungen über die Schutzscheide und die Bildung des Stammes und der Wurzel. *Jahrb. wissensch. Botanik* **4**, 101–124 (1865).
- Rufz de Lavison, J. d. Du mode de pénétration de quelques sels dans la plante vivante. *Revue générale Botanique* **22**, 225–241 (1910).
- Nagahashi, G., Thomson, W. W. & Leonard, R. T. The Casparian strip as a barrier to the movement of lanthanum in corn roots. *Science* **183**, 670–671 (1974).
- Steudle, E. Water uptake by plant roots: an integration of views. *Plant Soil* **226**, 45–56 (2000).
- van Fleet, D. S. Histochemistry and function of the endodermis. *Bot. Rev.* **27**, 165–220 (1961).
- Vasse, J., Frey, P. & Trigalet, A. Microscopic studies of intercellular infection and protoxylem invasion of tomato roots by *Pseudomonas solanacearum*. *Mol. Plant Microbe Interact.* **8**, 241–251 (1995).
- Heidstra, R., Welch, D. & Scheres, B. Mosaic analyses using marked activation and deletion clones dissect *Arabidopsis* SCARECROW action in asymmetric cell division. *Genes Dev.* **18**, 1964–1969 (2004).
- Carlsbecker, A. *et al.* Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* **465**, 316–321 (2010).
- Nakajima, K., Sena, G., Nawy, T. & Benfey, P. N. Intercellular movement of the putative transcription factor SHR in root patterning. *Nature* **413**, 307–311 (2001).
- Ma, J. F. *et al.* An efflux transporter of silicon in rice. *Nature* **448**, 209–212 (2007).
- Alassimone, J., Naseer, S. & Geldner, N. A developmental framework for endodermal differentiation and polarity. *Proc. Natl Acad. Sci. USA* **107**, 5214–5219 (2010).
- Takano, J. *et al.* Polar localization and degradation of *Arabidopsis* boron transporters through distinct trafficking pathways. *Proc. Natl Acad. Sci. USA* **107**, 5220–5225 (2010).
- Bonnett, H. T. Jr. The root endodermis: fine structure and function. *J. Cell Biol.* **37**, 199–205 (1968).
- Clarkson, D. T. Calcium transport between tissues and its distribution in the plant. *Plant Cell Environ.* **7**, 449–456 (1984).
- Birnbaum, K. *et al.* A gene expression map of the *Arabidopsis* root. *Science* **302**, 1956–1960 (2003).
- Di Laurenzio, L. *et al.* The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. *Cell* **86**, 423–433 (1996).
- Schreiber, L. Chemical composition of Casparian strips isolated from *Clivia miniata* Reg roots: evidence for lignin. *Planta* **199**, 596–601 (1996).
- Geldner, N. *et al.* The *Arabidopsis* GNOM ARF-GEF mediates endosomal recycling, auxin transport, and auxin-dependent plant growth. *Cell* **112**, 219–230 (2003).
- Karahara, I. & Shibaoka, H. Isolation of Casparian strips from pea roots. *Plant Cell Physiol.* **33**, 555–561 (1992).
- Reichardt, I. *et al.* Plant cytokinesis requires *de novo* secretory trafficking but not endocytosis. *Curr. Biol.* **17**, 2047–2053 (2007).
- Geldner, N. *et al.* Rapid, combinatorial analysis of membrane compartments in intact plants with a multicolor marker set. *Plant J.* **59**, 169–178 (2009).
- Sriranganadane, D. *et al.* *Aspergillus* protein degradation pathways with different secreted protease sets at neutral and acidic pH. *J. Proteome Res.* **9**, 3511–3519 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Central Imaging Facility (CIF) and the Protein Analysis Facility (PAF) of the University of Lausanne for technical support, G. Jürgens, W. Michalke, J. Chory for antibodies, T. W. J. Gadella for mTurquoise and J. Takano for mutant BOR1 clone. We thank J. Chory, E. M. N. Dohmann, Y. Lee and G. Jürgens for comments and critical reading of the manuscript. This work was supported by grants to N.G., the Belgian Interuniversity Attraction Poles program (IAP VI/33) and the Special Research Fund of Ghent University to B.D.R. and T.B., the Roche Research Foundation to D.R. and a Marie-Curie IEF grant to J.E.M.V.

Author Contributions N.G., D.R., B.D.R. and T.B. conceived and designed experiments. D.R., B.D.R., V.D.T., A.P., J.A., J.E.M.V., M.Y., Y.-D.S. and N.G. performed the experiments. D.R., B.D.R., V.D.T., A.P., J.A., J.E.M.V., Y.-D.S. and N.G. analysed the data. N.G. and D.R. wrote the manuscript.

Author Information Sequences have been deposited in GenBank under accession code HQ699533–50. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.G. (niko.geldner@unil.ch).

METHODS

Optical microscopy. Images were taken with a Leica SP/2 confocal microscope. Excitation and detection windows were set as follows: GFP 488 nm, 500–600 nm; mCherry 594 nm, 600–700 nm; GFP and FM4-64/propidium iodide 488 nm, 500–550 nm and 600–700 nm; Citrine/YFP and mCherry 514 nm and 594 nm, 520–560 nm and 600–700 nm. Images for CASP1–GFP three-dimensional reconstruction were taken with a 2-photon Zeiss LSM 710 NLO confocal, with a Chameleon Ultra II Ti:Sapphire laser at 820 nm using the non-descanned detector with a 500–550 bandpass filter. For autofluorescence of Casparian strips, excitation at 770 nm was used, with the same bandpass filter. BFA, propidium iodide and FM4-64 treatments were done as in ref. 13.

Electron microscopy. For ultrastructural analysis, roots of *A. thaliana* were cryo-fixed by high-pressure freezing (Baltec HPM 010) in hexadecane, freeze-substituted in acetone containing 2.5% osmium tetroxide and infiltrated with epon at 0 °C. Ultrathin resin sections were stained with 2% uranyl acetate in 50% ethanol and lead citrate²². For immunogold localization of GFP fusion proteins (CASP1–GFP and CASP5–GFP), roots were fixed with 4% formaldehyde and 8% formaldehyde for 45 min and 90–120 min, respectively. Root tips were embedded in 10% gelatine and infiltrated with a mixture of polyvinylpyrrolidone and sucrose²⁵. After freezing, ultrathin cryosections were sectioned using a Leica Ultracut UCT/EMFCS and the frozen sections were transferred to electron microscopic grids for immunogold labelling. The thawed sections were incubated with rabbit anti-GFP (1:500; Torrey Pines) and goat anti-rabbit F(ab)₂ fragments coupled to Nanogold (1:50; Nanoprobes). After silver enhancement (HQ Silver, 8 min; Nanoprobes) sections were embedded in methyl cellulose containing 0.45% uranyl acetate. Ultrathin resin and cryosections were viewed in a LEO 906 transmission electron microscope at 80 kV accelerating voltage. GFP–immunogold labelling on wild-type cells was negligible.

FRAP. Five-day-old seedlings were treated for 1 h with 50 mM 2-deoxyglucose and 0.02% sodium azide, and then imaged with a Leica SP/2 confocal microscope. Bleaching was performed on an area of 2 µm² by excitation at 488 nm for GFP, at 488 nm, 496 nm and 514 nm for citrine. Images were taken with the same settings before bleaching (time 0) and then every minute for 20 min. Fluorescence intensities were measured with ImageJ. Intensities were corrected by subtracting background and forming ratios using intensities from areas without signal (background) and non-bleached areas with signal (measuring overall bleaching). Finally, pre-bleaching and post-bleaching fluorescence were set at 100% and 0%, respectively, to allow averaging of individual experiments. After one pre-bleach scan, bleaching was done by one scan with 16× line average, 2× frame average, maximally zooming onto the bleach region of interest at maximal laser power. The first post-bleach scan was done at minute one after the pre-bleach scan. For a more detailed description of a similar protocol, see ref. 26.

Time-lapse imaging of CASP1 expression. For confocal time-lapse imaging in CASP1::CASP1–GFP, 5-day-old seedlings were transferred into a Lab-Tek II chambered coverglass (Nunc) and covered with a small block of agar to prevent drying. Subsequently, slides were mounted and imaged on an inverted Leica SP/2 confocal microscope. Image stacks were taken every 3 min for a total time of 2 h. Confocal images were analysed and processed using the IMAGEJ plugin of the MBF IMAGEJ bundle.

Protein fractionation. One-hundred milligrams of roots from 5-day-old seedlings were frozen in liquid nitrogen and ground using Qiagen Tissue-lyser.

Two-hundred microlitres of extraction buffer (50 mM HEPES pH 7.9, 300 mM sucrose, 150 mM NaCl, 5 mM EDTA, 10 mM potassium acetate, 2.5× Roche Complete Protein inhibitors, 1 mM PMSF) were then added and samples were spun at 6,500g, 4 °C for 5 min. Supernatant (S_{6,500}) was collected and pellets washed three times with extraction buffer. The last pellet was treated with extraction buffer containing additional detergents (see figure legends for concentrations).

Immunoprecipitation/MS analysis. All experiments were done using tagged CASP variants expressed under their endogenous promoters. P_{6,500} from 450 mg of roots was treated for 30 min on ice with 450 µl extraction buffer plus 1% CHAPS, and then spun at 19,000g, 4 °C for 30 min. 410 µl were incubated on ice for 30 min with 50 µl of µMACS anti-GFP microbeads. Beads were then washed on column with 4× 200 µl and 1× 100 µl extraction buffer plus 1% CHAPS. Elution was performed at 95 °C with 50 µl 1× µMACS elution buffer. For western blots, 15 µl (3.3% of input, 30% of output) were charged on gel. GFP was probed with a rabbit polyclonal anti-GFP (Invitrogen, 1:1,000), mCherry with a rabbit polyclonal anti-DsRed (Clontech, 1:1,000). An HRP-coupled anti-rabbit secondary antibody (Promega) was used at 1:1,000 for GFP, 1:20,000 for mCherry. For IP-MS, 2 g of roots were used; P_{6,500} was re-suspended in extraction buffer + 0.5% Triton X-100. After gel migration and trypsinization, samples were analysed on a hybrid linear trap LTQ-Orbitrap mass spectrometer (Thermo Fisher) interfaced to an Agilent 1100 nano HPLC system. Peptides were loaded onto a trapping microcolumn ZORBAX 300SB C18 (Agilent), eluted after 5 min and separated on a reversed-phase nanocolumn ZORBAX 300SB C18 column (Agilent). A 400 nozzle ESI Chip (Advion Biosciences) was used for spraying, with a voltage of 1.65 kV and mass spectrometer capillary transfer temperature of 200 °C. In data-dependent acquisition controlled by Xcalibur 2.0.7 software (Thermo Fisher), the six most intense precursor ions detected in the full MS survey were selected and fragmented. All MS/MS samples were analysed using Mascot 2.2 searching UNIPROT database. Scaffold_3 was used to validate MS/MS-based peptide and protein identifications, and to perform data set alignment. 90.0% and 95% probability thresholds were used to identify peptide and proteins, respectively.

Transgenic lines. Transgenic lines were generated by floral dipping of *A. thaliana*, ecotype Col-0 (ref. 27). Sequences of plasmids used to generate transgenic lines are deposited in GenBank (HQ699533–50). RNA interference was done using the following clones from the AGRICOLA project: CASP1 CATMA2a34330 (nucleotides 2–231 of the CDS); CASP2 CATMA3a10500 (nucleotides 595–754 of the CDS and 1–139 of the 3' UTR); CASP3 CATMA2a25770 (nucleotides 2–161 of the CDS); CASP5 CATMA5a13570 (nucleotides 3–162 of the CDS). Artificial microRNAs were designed using the Web microRNA Designer (<http://wmd2.weigelworld.org>) and expressed under the CASP1 promoter. The following positions were targeted: CASP1 nucleotides 547–567 of the CDS; CASP2 nucleotides 66–86 of the 3' UTR; CASP3 nucleotides 624–644 of the CDS; CASP4 nucleotides 277–297 of the CDS; CASP5 nucleotides 150–170 of the CDS.

25. Tokuyasu, K. T. Use of poly(vinylpyrrolidone) and poly(vinyl alcohol) for cryoultramicrotomy. *Histochem. J.* **21**, 163–171 (1989).
26. Men, S. *et al.* Sterol-dependent endocytosis mediates post-cytokinetic acquisition of PIN2 auxin efflux carrier polarity. *Nature Cell Biol.* **10**, 237–244 (2008).
27. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).

BCL6 enables Ph⁺ acute lymphoblastic leukaemia cells to survive BCR–ABL1 kinase inhibition

Cihangir Duy^{1,2}, Christian Hurtz^{1,6}, Seyedmehdi Shojaei¹, Leandro Cerchietti³, Huimin Geng³, Srividya Swaminathan¹, Lars Klemm¹, Soo-mi Kweon², Rahul Nahar^{1,2}, Melanie Braig⁴, Eugene Park², Yong-mi Kim², Wolf-Karsten Hofmann⁵, Sebastian Herzog⁶, Hassan Jumaa⁶, H. Phillip Koeffler⁷, J. Jessica Yu⁸, Nora Heisterkamp², Thomas G. Graeber⁹, Hong Wu⁹, B. Hilda Ye⁸, Ari Melnick³ & Markus Müschen^{1,2}

Tyrosine kinase inhibitors (TKIs) are widely used to treat patients with leukaemia driven by BCR–ABL1 (ref. 1) and other oncogenic tyrosine kinases^{2,3}. Recent efforts have focused on developing more potent TKIs that also inhibit mutant tyrosine kinases^{4,5}. However, even effective TKIs typically fail to eradicate leukaemia-initiating cells (LICs)^{6–8}, which often cause recurrence of leukaemia after initially successful treatment. Here we report the discovery of a novel mechanism of drug resistance, which is based on protective feedback signalling of leukaemia cells in response to treatment with TKI. We identify BCL6 as a central component of this drug-resistance pathway and demonstrate that targeted inhibition of BCL6 leads to eradication of drug-resistant and leukaemia-initiating subclones.

BCL6 is a known proto-oncogene that is often translocated in diffuse large B-cell lymphoma (DLBCL)⁹. In response to TKI treatment, BCR–ABL1 acute lymphoblastic leukaemia (ALL) cells upregulate BCL6 protein levels by approximately 90-fold: that is, to similar levels as in DLBCL (Fig. 1a). Upregulation of BCL6 in response to TKI treatment represents a novel defence mechanism, which enables leukaemia cells to survive TKI treatment: Previous work suggested that TKI-mediated cell death is largely p53 independent. Here we demonstrate that BCL6 upregulation upon TKI treatment leads to transcriptional inactivation of the p53 pathway. BCL6-deficient leukaemia cells fail to inactivate p53 and are particularly sensitive to TKI treatment. BCL6^{−/−} leukaemia cells are poised to undergo cellular senescence and fail to initiate leukaemia in serial transplant recipients. A combination of TKI treatment and a novel BCL6 peptide inhibitor markedly increased survival of NOD/SCID mice xenografted with patient-derived BCR–ABL1 ALL cells. We propose that dual targeting of oncogenic tyrosine kinases and BCL6-dependent feedback (Supplementary Fig. 1) represents a novel strategy to eradicate drug-resistant and leukaemia-initiating subclones in tyrosine-kinase-driven leukaemia.

To elucidate mechanisms of TKI resistance in tyrosine-kinase-driven leukaemia, we performed a gene expression analysis including our and published data of TKI-treated leukaemia. We identified BCL6 as a top-ranking gene in a set of recurrent gene expression changes, some of which are shared with mitogen-activated protein-kinase (MEK) inhibition in BRAF^{V600E} mutant solid tumour cells¹⁰ (Supplementary Figs 2 and 3). TKI-induced upregulation of BCL6 messenger RNA (mRNA) levels was confirmed in multiple leukaemia subtypes carrying oncogenic tyrosine kinases (Supplementary Fig. 2). The BCR–ABL1 kinase, encoded by the Philadelphia chromosome (Ph), represents the most frequent genetic lesion in adult ALL, defines the subtype with a particularly poor prognosis^{1,4,5} and was therefore chosen as focus for this study.

To elucidate the regulation of BCL6 in Ph⁺ ALL, we investigated the JAK2/STAT5 (ref. 11) and PI3K/AKT¹² pathways downstream of BCR–ABL1. We and others have shown that STAT5 suppresses BCL6 in B cells^{13–15}. TKI-mediated upregulation of BCL6 was diminished by constitutively active STAT5 (Fig. 1b) and deletion of STAT5 was sufficient to upregulate BCL6, even in the absence of TKI treatment (Fig. 1c). In agreement with previous work¹⁵, overexpression of FoxO4 induced BCL6 (Fig. 1d). In Ph⁺ ALL cells, FoxO factors are inactivated by PI3K/AKT¹² signalling, which is reversed by Pten (Supplementary Fig. 4). Deletion of *Pten*, hence, abrogated the ability of the leukaemia cells to upregulate BCL6 in response to TKI treatment (Fig. 1e).

In DLBCL, BCL6 is frequently translocated and suppresses p53-mediated apoptosis^{9,16}. Although TKI treatment is less effective in p53^{−/−} Ph⁺ ALL¹⁷, recent studies showed that TKI paradoxically prevents the upregulation of p53 in response to DNA damage in Ph⁺ ALL and chronic myelogenous leukaemia^{18,19}. A comparative gene expression analysis of BCL6^{−/−} and BCL6^{+/+} ALL cells (Supplementary Fig. 5) identified Cdkn2a (Arf), Cdkn1a (p21), p53 and p53bp1 as potential BCL6 target genes (Supplementary Fig. 6). Arf and p53 protein levels were indeed unrestrained in BCL6^{−/−} ALL (Fig. 2a). TKI treatment of BCL6^{+/+} ALL resulted in strong upregulation of BCL6 with low levels of p53, whereas BCL6^{−/−} ALL cells failed to curb p53 expression levels (Supplementary Fig. 7). Likewise, TKI treatment increased excessively p53 levels when *Pten*-deficient ALL cells failed to upregulate BCL6 (Fig. 1e).

Identifying direct targets of BCL6 by chromatin immunoprecipitation (ChIP) in Ph⁺ ALL (Supplementary Figs 8–11), p53, p21 and p27 were among the genes with the strongest recruitment of BCL6 in TKI-treated ALL (Fig. 2b and Supplementary Figs 9–11). Given that cell-cycle arrest and senescence-associated genes were among the BCL6 targets, we studied the cell-cycle profile of leukaemia cells. BCL6^{−/−} ALL cells divided at a slightly reduced rate compared with BCL6^{+/+} ALL cells (Fig. 2c). Treatment with adriamycin (0.05 µg ml^{−1}) had no significant effect on BCL6^{+/+} ALL cells in a senescence-associated β-galactosidase assay^{20,21} but revealed that most BCL6^{−/−} leukaemia cells were poised to undergo cellular senescence (Fig. 2d). These findings demonstrate that even low levels of BCL6 in the absence of TKI treatment are critical to downregulate Arf/p53.

Clonal evolution of leukaemia involves acquisition of genetic lesions through DNA damage²². Interestingly, a comparative genomic hybridization analysis revealed that genetic lesions were less frequent in BCL6 deficient ALL (Supplementary Fig. 12), suggesting that increased sensitivity to DNA damage limits clonal evolution in the absence of BCL6. Because Arf and p53 are critical negative regulators of self-renewal²³, we performed colony-forming assays. The colony frequencies of

¹Department of Laboratory Medicine, University of California San Francisco, San Francisco, California 94143, USA. ²Children's Hospital Los Angeles, Department of Pediatrics, University of Southern California, Los Angeles, California 90027, USA. ³Departments of Medicine and Pharmacology, Weill Cornell Medical College, New York, New York 10065, USA. ⁴Department of Hematology and Oncology, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany. ⁵Department of Hematology and Oncology, Universität Heidelberg, Klinikum Mannheim, Mannheim, Germany. ⁶Faculty of Biology, BIOS Centre for Biological Signalling Studies, Albert-Ludwigs-Universität Freiburg and Max-Planck-Institute for Immunobiology, Freiburg, Germany. ⁷Division of Hematology and Oncology, Cedars Sinai Medical Center, Los Angeles, California 90095, USA. ⁸Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ⁹Department of Molecular Pharmacology, University of California Los Angeles, Los Angeles, California 90095, USA.

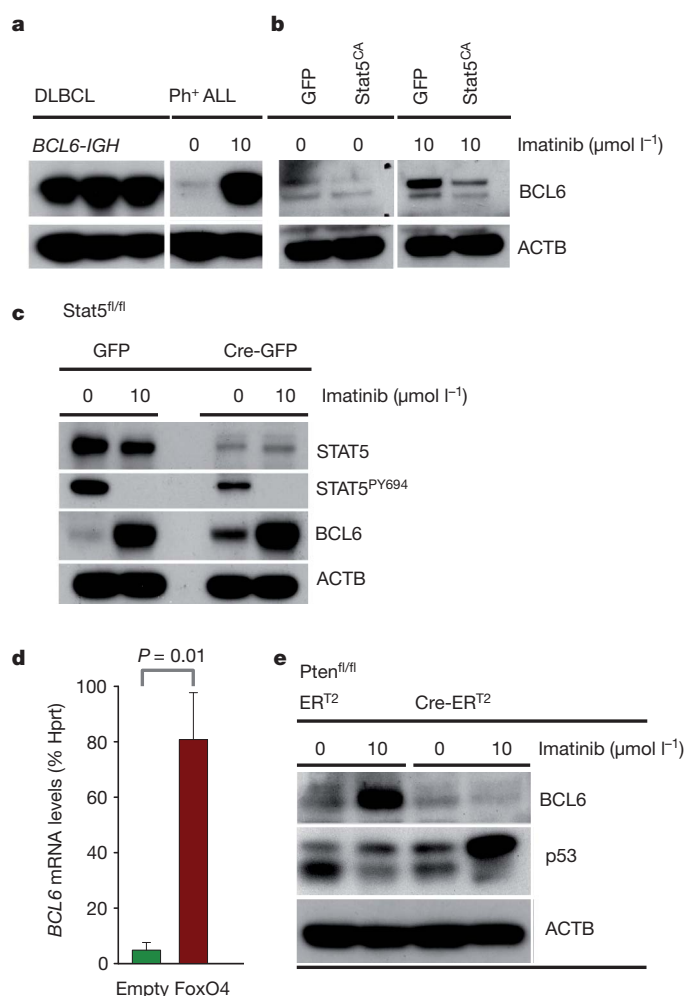


Figure 1 | Regulation of BCL6 expression in BCR-ABL1 ALL cells. **a**, Ph⁺ ALL cells were treated with and without imatinib ($10 \mu\text{mol l}^{-1}$) for 24 h. Upregulation of BCL6 was compared with expression levels in DLBCL by western blot. **b**, BCR-ABL1-transformed mouse ALL cells were transduced with a constitutively active Stat5 mutant (STAT5^{CA}) or a control vector (green fluorescent protein, GFP) and treated either with or without imatinib. BCL6 western blot was performed using β -actin as loading control. **c**, BCL6 expression upon imatinib treatment was studied by western blot in the presence or absence of Cre-mediated deletion of *Stat5* in BCR-ABL1-transformed Stat5^{fl/fl} mouse ALL. **d**, Mouse BCR-ABL1 ALL cells were transduced with FoxO4-puromycin or a puromycin control vector and subjected to antibiotic selection. Cells were collected and BCL6 mRNA levels were measured by qRT-PCR relative to Hprt. **e**, Imatinib-induced BCL6 expression was studied by western blot in the presence or absence of Cre-mediated deletion of *Pten* in BCR-ABL1-transformed Pten^{fl/fl} mouse ALL cells.

BCL6^{-/-} ALL cells were reduced by approximately 20-fold compared with BCL6^{+/+} ALL (Fig. 3a). To study self-renewal *in vivo*, we measured the ability of BCL6^{+/+} and BCL6^{-/-} ALL cells to initiate leukaemia in transplant recipients (Fig. 3b). Using luciferase bioimaging, leukaemia engraftment was observed in both groups after 8 days. BCL6^{+/+} ALL cells rapidly expanded and initiated fatal leukaemia, whereas BCL6^{-/-} ALL cells failed to expand from the initial engraftment foci (Fig. 3c). Some mice that received BCL6^{-/-} ALL cells ultimately succumbed to leukaemia (Fig. 3b). Flow cytometry, however, revealed that the leukaemias in the BCL6^{-/-} group were in fact derived from endogenous CD45.1⁺ cells of the irradiated recipients and not from the injected CD45.2⁺ donor ALL cells (Supplementary Fig. 13 and asterisks in Fig. 3b).

Defective leukaemia initiation may be a consequence of impaired homing to the bone marrow niche. Indeed, BCL6^{-/-} ALL cells lack

expression of CD44 (Supplementary Fig. 14), which is critical for homing of BCR-ABL1 LICs to the bone marrow microenvironment²⁴. Retroviral reconstitution of CD44 markedly increased homing of BCL6^{-/-} ALL cells to the bone marrow niche, but failed to rescue defective leukaemia initiation (Supplementary Fig. 14).

Using intrafemoral injection to circumvent homing defects, a limiting dilution experiment (Fig. 3d) showed that 5 million BCL6^{-/-} ALL cells compared with only 10^3 BCL6^{+/+} ALL cells were needed to initiate fatal leukaemia. These findings suggest that the frequency of LIC in BCL6^{-/-} ALL (fewer than 1 in 100,000) is reduced by more than 100-fold compared with BCL6^{+/+} ALL (at least 1 in 1,000). An alternative interpretation would be that LICs occur at a similar frequency in BCL6^{-/-} ALL but with reduced self-renewal activity. To address potential 'exhaustion' of LICs, we performed a serial transplantation with ALL cells that gave rise to disease in primary recipients after injection of 5 million ALL cells. From the bone marrow, we isolated CD19⁺ ALL cells for secondary intrafemoral injection. BCL6^{-/-} leukaemia was not transplantable in secondary recipients (Supplementary Fig. 15). Although these findings do not exclude the possibility that the LIC frequencies are reduced in BCL6^{-/-} ALL, they support the notion of LIC 'exhaustion' after secondary transplantation.

To explore the therapeutic usefulness of pharmacological inhibition of BCL6, we tested a BCL6 inhibitor (*retro-inverso* BCL6 peptide-inhibitor (RI-BPI)), which blocks the repressor activity of BCL6 (ref. 25). Gene expression analysis confirmed that RI-BPI is a selective and potent inhibitor of BCL6 (Supplementary Fig. 16). We investigated the effect of RI-BPI on the self-renewal capacity of primary Ph⁺ ALL and the initiation of leukaemia in a mouse xenograft model. Treatment with RI-BPI resulted in a reduction of colony formation and delayed progression of leukaemia. Likewise, treatment of Ph⁺ ALL with RI-BPI induced cellular senescence (Supplementary Fig. 17).

We next examined how gene dosage of BCL6 affects responses to TKI. For instance, Pten^{-/-} ALL cells lack the ability to upregulate the p53-repressor BCL6 and are more sensitive to imatinib (Fig. 1e and Supplementary Fig. 18). Dose-response studies in BCL6^{+/+}, BCL6^{+/-} and BCL6^{-/-} ALL (Fig. 4a) showed that sensitivity to imatinib was significantly increased in BCL6^{-/-} (half maximal effective concentration (EC₅₀) $0.17 \mu\text{mol l}^{-1}$) and even in BCL6^{+/-} ALL cells (EC₅₀ $0.67 \mu\text{mol l}^{-1}$) compared with BCL6^{+/+} ALL cells (EC₅₀ $1.10 \mu\text{mol l}^{-1}$; Fig. 4a). These findings indicate that maximum levels of BCL6 expression are required to prevent TKI-induced cell death. Indeed, inducible activation of BCL6-ERT² constructs²⁶ in BCL6^{-/-} ALL cells conferred a strong survival advantage in the presence of imatinib (Fig. 4b). Activation of BCL6 in BCL6^{+/+} ALL cells induced cell-cycle exit (not shown) and no additional survival advantage, because these cells already achieved maximal upregulation of endogenous BCL6 (Fig. 1a).

To address the role of BCL6-mediated repression of p53 in TKI-resistance, p53^{-/-} and p53^{+/+} ALL cells were treated with RI-BPI. The synergistic effect between TKI treatment and RI-BPI is indeed partly p53 dependent (Supplementary Fig. 19). In p53^{-/-} ALL cells, the effect of RI-BPI was significantly diminished compared with p53^{+/+} ALL.

To confirm that BCL6 has a similar function in patient-derived Ph⁺ ALL, primary ALL cells were transduced with a dominant-negative BCL6 mutant (DN-BCL6-ERT²)²⁶, which resulted in a marked competitive disadvantage of Ph⁺ ALL cells, that was further enhanced by imatinib treatment (Fig. 4c). Similar observations in mouse ALL and in an established Ph⁺ ALL cell line demonstrated that BCL6 promotes survival of TKI-treated Ph⁺ ALL (Supplementary Fig. 20).

To test the effect of BCL6 inhibition on TKI resistance, we cultured four primary Ph⁺ ALL in the presence or absence of imatinib, RI-BPI or a combination of both (Supplementary Fig. 21). Initially, all four Ph⁺ ALL cases responded to imatinib treatment, but subsequently rebounded and were no longer sensitive to imatinib ($10 \mu\text{mol l}^{-1}$). RI-BPI alone showed only slight effects, whereas the combination of RI-BPI and imatinib rapidly induced cell death and effectively prevented a rebound in all four cases (Supplementary Fig. 21). These

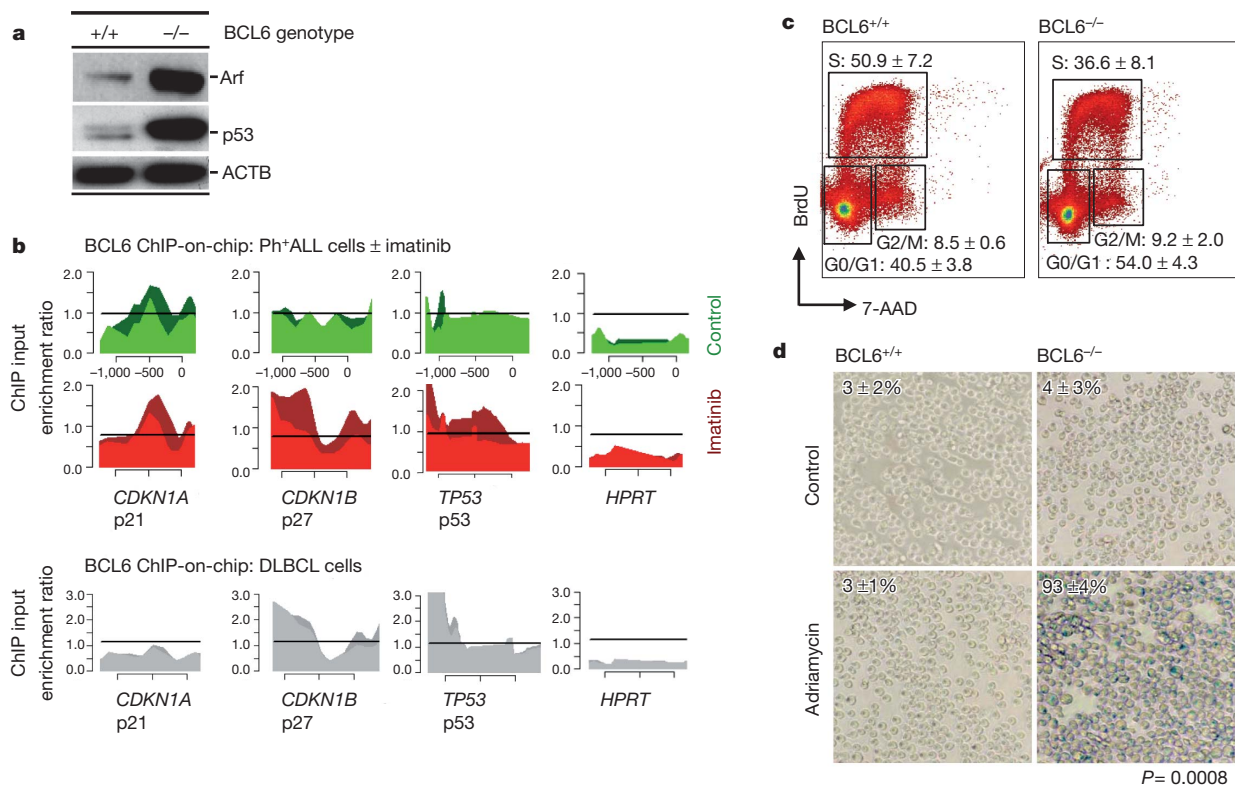


Figure 2 | BCL6 is required for transcriptional inactivation of the Arf/p53 pathway in BCR-ABL1 ALL. **a**, Western blot analysis of CDKN2A (Arf) and p53 expression in BCL6^{-/-} and BCL6^{+/+} BCR-ABL1 ALL cells. **b**, Human Ph⁺ ALL cells (Tom1) were treated with and without imatinib ($10 \mu\text{mol l}^{-1}$) for 24 h and were subjected to ChIP-on-chip analysis using a BCL6-specific antibody. The y axis indicates enrichment versus input, the x axis the location of probes within the respective loci relative to the transcriptional start site. The

findings suggest that prolonged treatment with a combination of imatinib/RI-BPI prevents acquisition of TKI-resistance. We next examined the effect of imatinib/RI-BPI combinations on primary TKI-resistance in Ph⁺ ALL. To this end, four human Ph⁺ ALL cell lines that lacked BCR-ABL1 kinase mutations (Supplementary Table 1)

dark and light green (control) or red (imatinib) tracings depict two replicates. Recruitment to *CDKN1A*, *CDKN1B*, *TP53* and *HPRT* (negative control) is shown in Ph⁺ ALL cells and one DLBCL cell line (OCI-Ly7). **c**, Cell-cycle analysis (BrdU/7-AAD staining). **d**, Staining for senescence-associated β -galactosidase (SA- β -gal). ALL cells were treated with or without $0.05 \mu\text{g ml}^{-1}$ adriamycin for 48 h to induce a low level of DNA damage. Percentages of SA- β -gal⁺ cells are indicated (means \pm SD; $n = 3$).

but which were highly refractory to imatinib ($10 \mu\text{mol l}^{-1}$) were treated with or without imatinib, RI-BPI or a combination of both. Imatinib alone did not achieve a therapeutic response, whereas the combination with RI-BPI potentiated the effect of imatinib on the refractory ALL cells (Supplementary Fig. 22).

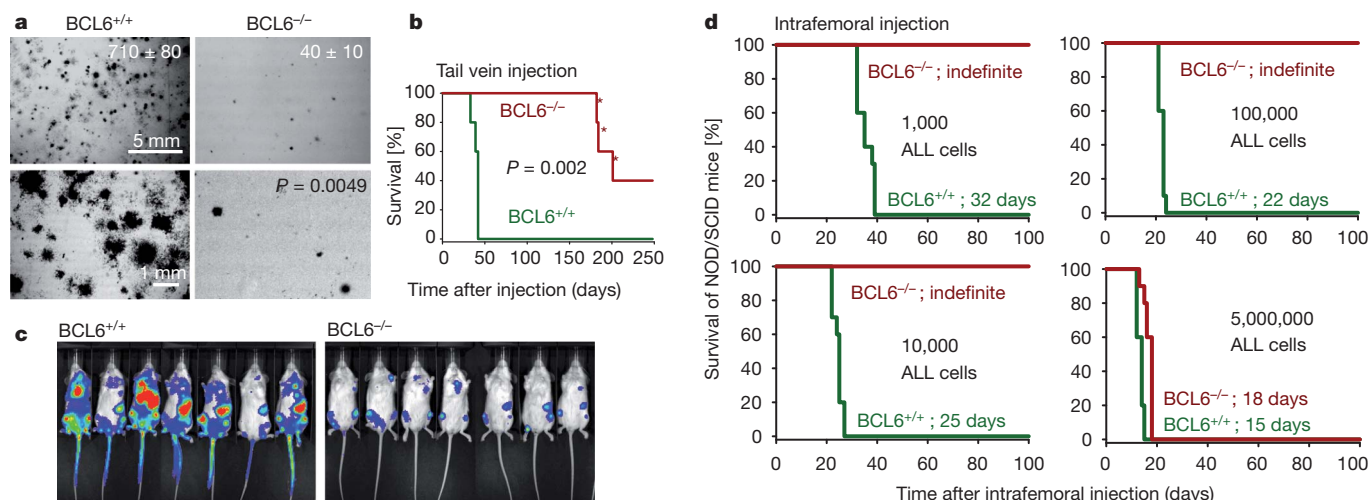


Figure 3 | BCL6 is required for leukaemia initiation in BCR-ABL1 ALL. **a**, Ten thousand BCL6^{-/-} or BCL6^{+/+} BCR-ABL1 ALL cells were plated in semisolid agar, and colonies were counted after 10 days (numbers denote means \pm SD, $n = 3$). **b**, Overall survival of mice injected with 100,000 BCL6^{-/-} and BCL6^{+/+} BCR-ABL1 ALL cells was compared by Kaplan-Meier analysis. Mice that developed CD45.1⁺ endogenous leukaemia instead of leukaemia from injected CD45.2⁺ cells are indicated by asterisks (see Supplementary

Fig. 13). **c**, For an SCID LIC (SL-IC) experiment, BCL6^{-/-} and BCL6^{+/+} BCR-ABL1 ALL cells were labelled with firefly luciferase and intravenously injected into sublethally irradiated NOD/SCID mice. **d**, The SL-IC assay was repeated as a limiting dilution experiment (10^3 , 10^4 , 10^5 , 5 million cells) and leukaemia cells were directly injected into the femoral bone marrow to circumvent potential engraftment defects.

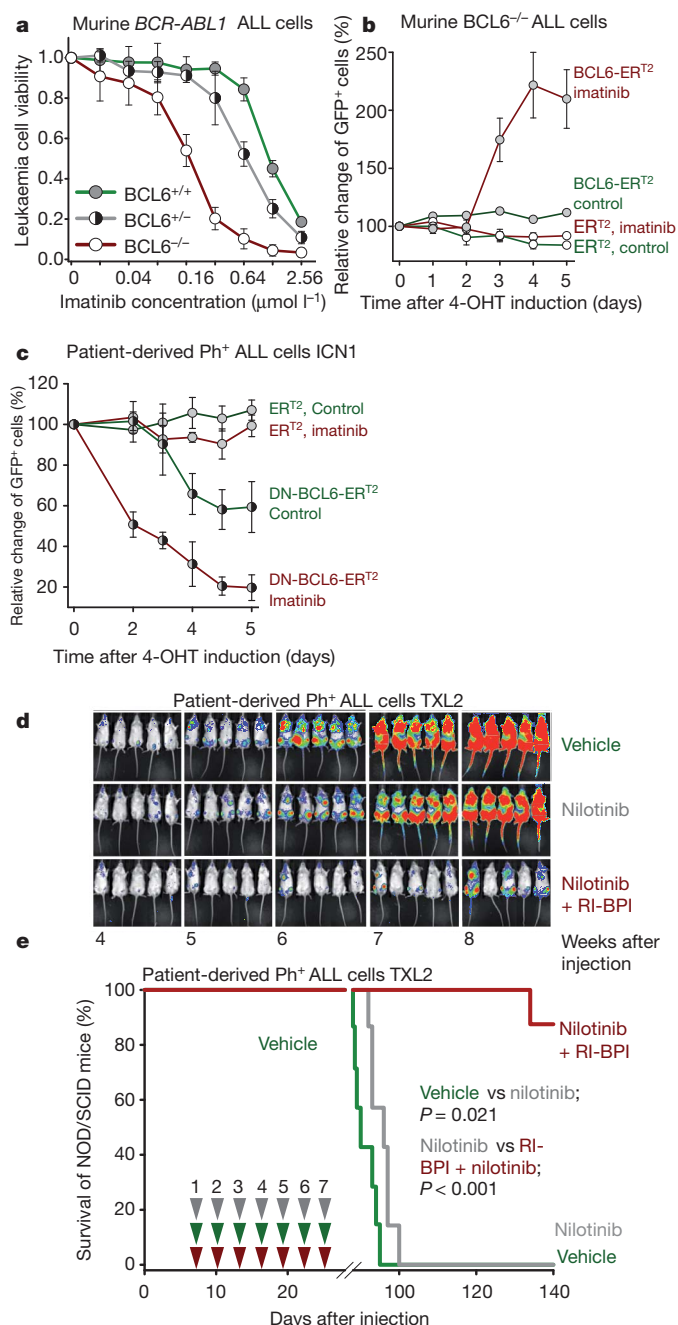


Figure 4 | **BCL6 promotes survival of TKI-treated BCR-ABL1 ALL cells.**

a, Imatinib sensitivity of *BCL6*^{-/-}, *BCL6*^{+/-} and *BCL6*^{+/+} ALL cells was measured in a resazurin viability assay. **b**, *BCL6*^{-/-} ALL cells were transduced with *BCL6-ER*^{T2} or *ER*^{T2} vectors (tagged with GFP). ALL cells were treated with or without $1 \mu\text{mol l}^{-1}$ imatinib, and *BCL6-ER*^{T2} or *ER*^{T2} were induced by 4-hydroxytamoxifen. Relative changes of GFP⁺ cells after induction are indicated. **c**, Patient-derived Ph⁺ ALL cells (ICN1) were transduced with inducible dominant-negative BCL6 (DN-*BCL6-ER*^{T2}) or *ER*^{T2} control vectors. ALL cells were treated with or without $10 \mu\text{mol l}^{-1}$ imatinib and DN-*BCL6-ER*^{T2} or *ER*^{T2} were induced by 4-hydroxytamoxifen. Relative changes of GFP⁺ cells after induction are indicated. **d**, Patient-derived Ph⁺ ALL cells (TXL2) were labelled with luciferase and 100,000 cells were injected. Mice were treated seven times with either vehicle (green), nilotinib (25 mg kg^{-1} ; grey) or a combination of nilotinib and RI-BPI (25 mg kg^{-1} ; red). Treated mice are shown in **e**, a Kaplan-Meier survival analysis. Treatment days are indicated by arrowheads.

To study the efficacy of combined tyrosine kinase and BCL6 inhibition *in vivo*, primary Ph⁺ ALL cells were labelled with luciferase and xenografted into mice. Recipient mice were treated with either vehicle, nilotinib or a combination of nilotinib and RI-BPI. Nilotinib is more potent than imatinib, which only has marginal effects in mice^{27,28}. Bioimaging demonstrated that seven to ten injections of RI-BPI significantly enhanced the effect of nilotinib (Fig. 4d, e and Supplementary Fig. 23). Whereas all mice treated with nilotinib alone succumbed to leukaemia within 99 days after injection, seven of eight mice treated with RI-BPI/nilotinib combination were still alive after 140 days (Fig. 4d, e). Also, in a model for full-blown mouse leukaemia, TKI/RI-BPI combinations proved effective and significantly prolonged survival (Supplementary Fig. 24). Establishing a potential therapeutic window of nilotinib/RI-BPI combinations, we found no evidence of relevant toxicity (Supplementary Figs 25 and 26 and Supplementary Table 2).

Although transcription factors have been considered intractable therapeutic targets, the recent development of a small molecule inhibitor against BCL6 (ref. 29) holds promise for effectively targeting TKI-resistance in patients with Ph⁺ ALL. Because TKI-resistance develops in virtually all cases of Ph⁺ ALL, it appears particularly important to target this novel pathway of TKI-resistance.

METHODS SUMMARY

Cell culture. Primary leukaemia cells (Supplementary Table 1) were cultured on OP9 stroma cells in alpha minimum essential medium without ribonucleotides and deoxyribonucleotides, supplemented with 20% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 100 IU ml⁻¹ penicillin and 100 $\mu\text{g ml}^{-1}$ streptomycin. Human ALL cell lines were maintained in RPMI with GlutaMAX containing 20% FBS, 100 IU ml⁻¹ penicillin and 100 $\mu\text{g ml}^{-1}$ streptomycin. Mouse *BCR-ABL1*-transformed ALL cells were maintained in IMDM with GlutaMAX containing 20% FBS, 100 IU ml⁻¹ penicillin, 100 $\mu\text{g ml}^{-1}$ streptomycin and 50 μM 2-mercaptoethanol. Cell cultures were kept at 37 °C in a humidified incubator under a 5% CO₂ atmosphere.

BCR-ABL1 transfection. Transfection of a murine stem cell virus (MSCV)-based retroviral vector encoding BCR-ABL1 was performed using Lipofectamine 2000. Retroviral supernatant was produced by co-transfecting 293FT cells with the plasmids pHIT60 and pHIT123. Virus supernatant was collected, filtered through a 0.45 μm filter and loaded by centrifugation (2,000g, 90 min at 32 °C) on 50 $\mu\text{g ml}^{-1}$ RetroNectin-coated non-tissue well plates. Extracted bone marrow cells from mice were transduced by BCR-ABL1 in the presence of 10 ng ml⁻¹ recombinant murine interleukin-7 in RetroNectin-coated Petri dishes.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 February 2010; accepted 31 January 2011.

1. Druker, B. J. *et al.* Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *N. Engl. J. Med.* **344**, 1038–1042 (2001).
2. Armstrong, S. A. *et al.* Inhibition of FLT3 in MLL. Validation of a therapeutic target identified by gene expression based classification. *Cancer Cell* **3**, 173–183 (2003).
3. Meydan, N. *et al.* Inhibition of acute lymphoblastic leukaemia by a Jak-2 inhibitor. *Nature* **379**, 645–648 (1996).
4. Shah, N. P. *et al.* Overriding imatinib resistance with a novel ABL kinase inhibitor. *Science* **305**, 399–401 (2004).
5. O'Hare, T. *et al.* AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. *Cancer Cell* **16**, 401–412 (2009).
6. Graham, S. M. *et al.* Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to STI571 *in vitro*. *Blood* **99**, 319–325 (2002).
7. Naka, K. *et al.* TGF- β -FOXO signalling maintains leukaemia-initiating cells in chronic myeloid leukaemia. *Nature* **463**, 676–680 (2010).
8. Oravec-Wilson, K. I. *et al.* Persistence of leukemia-initiating cells in a conditional knockin model of an imatinib-responsive myeloproliferative disorder. *Cancer Cell* **16**, 137–148 (2009).
9. Saito, M. *et al.* A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell* **12**, 280–292 (2007).
10. Pratilas, C. A. *et al.* (V600E)BRAF is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. *Proc. Natl Acad. Sci. USA* **106**, 4519–4524 (2009).

11. Choudhary, C. *et al.* Mislocalized activation of oncogenic RTKs switches downstream signaling outcomes. *Mol. Cell* **36**, 326–339 (2009).
12. Janes, M. R. *et al.* Effective and selective targeting of leukemia cells using a TORC1/2 kinase inhibitor. *Nature Med.* **16**, 205–213 (2010).
13. Walker, S. R., Nelson, E. A. & Frank, D. A. STAT5 represses BCL6 expression by binding to a regulatory region frequently mutated in lymphomas. *Oncogene* **26**, 224–233 (2007).
14. Duy, C. *et al.* BCL6 is critical for the development of a diverse primary B cell repertoire. *J. Exp. Med.* **207**, 1209–1221 (2010).
15. Fernandez de Mattos, S. *et al.* FoxO3a and BCR-ABL regulate cyclin D2 transcription through a STAT5/BCL6-dependent mechanism. *Mol. Cell. Biol.* **24**, 10058–10071 (2004).
16. Phan, R. T. & Dalla-Favera, R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature* **432**, 635–639 (2004).
17. Wendel, H. G. *et al.* Loss of p53 impedes the antileukemic response to BCR-ABL inhibition. *Proc. Natl Acad. Sci. USA* **103**, 7444–7449 (2006).
18. Goldberg, Z., Levav, Y., Krichevsky, S., Fibach, E. & Haupt, Y. Treatment of chronic myeloid leukemia cells with imatinib (ST1571) impairs p53 accumulation in response to DNA damage. *Cell Cycle* **3**, 1188–1195 (2004).
19. Skorta, I. *et al.* Imatinib mesylate induces cisplatin hypersensitivity in Bcr-Abl+ cells by differential modulation of p53 transcriptional and proapoptotic activity. *Cancer Res.* **69**, 9337–9345 (2009).
20. Kamijo, T. *et al.* Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell* **91**, 649–659 (1997).
21. Braig, M. *et al.* Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* **436**, 660–665 (2005).
22. Mullighan, C. G. *et al.* Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377–1380 (2008).
23. Williams, R. T., Roussel, M. F. & Sherr, C. J. Arf gene loss enhances oncogenicity and limits imatinib response in mouse models of Bcr-Abl-induced acute lymphoblastic leukemia. *Proc. Natl Acad. Sci. USA* **103**, 6688–6693 (2006).
24. Krause, D. S., Lazarides, K., von Andrian, U. H. & Van Etten, R. A. Requirement for CD44 in homing and engraftment of BCR-ABL-expressing leukemic stem cells. *Nature Med.* **12**, 1175–1180 (2006).
25. Cerchietti, L. C. *et al.* A peptomimetic inhibitor of BCL6 with potent antilymphoma effects *in vitro* and *in vivo*. *Blood* **113**, 3397–3405 (2009).
26. Shaffer, A. L. *et al.* BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity* **13**, 199–212 (2000).
27. Williams, R. T. den, B. W. & Sherr, C. J. Cytokine-dependent imatinib resistance in mouse BCR-ABL+, Arf-null lymphoblastic leukemia. *Genes Dev.* **21**, 2283–2287 (2007).
28. Gruber, T. A., Chang, M. S., Spoto, R. & Muschen, M. Activation-induced cytidine deaminase accelerates clonal evolution in BCR-ABL1-driven B-cell lineage acute lymphoblastic leukemia. *Cancer Res.* **70**, 7411–7420 (2010).
29. Cerchietti, L. C. *et al.* A small-molecule inhibitor of BCL6 kills DLBCL cells *in vitro* and *in vivo*. *Cancer Cell* **17**, 400–411 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Dalla-Favera and L. Hennighausen for sharing BCL6^{-/-} and STAT5^{fl/fl} mice and wild-type controls with us. We thank A. L. Shaffer and L. M. Staudt for sharing their inducible BCL6 constructs. This work was supported by grants from the National Institutes of Health/National Cancer Institute through R01CA104348 (to A.M.), R01CA085573 (to B.H.Y.), R01CA026038 (to H.P.K.), R01CA090321 (to N.H.), R01CA137060 (to M.M.), R01CA139032 (to M.M.), R01CA157664 (to M.M.) and R21CA152497 (to M.M.), grants from the Leukemia and Lymphoma Society (to M.M.) Leukemia and Lymphoma Society SCOR 7005-11 (PI B. J. Druker), a grant from the Alex's Lemonade Stand Foundation for Pediatric Cancer Research (to M.M.), the California Institute for Regenerative Medicine through TR02-1816 (to M.M.), the William Laurence and Blanche Hughes Foundation and a Stand Up To Cancer-American Association for Cancer Research Innovative Research Grant IRG00909 (to M.M.). A.M. and M.M. are Scholars of the Leukemia and Lymphoma Society.

Author Contributions C.D. and M.M. conceived the study and wrote the paper. M.M. and A.M. designed experiments and interpreted data. C.D., C.H., S. Shojaaee, L.C., S. Swaminathan, L.K., S.-m.K., R.N., M.B., E.P. and Y.-m.K. designed and performed experiments and interpreted data. W.-K.H., H.P.K. and N.H. provided and characterized samples from patients. H.G. and T.G.G. analysed data. S.H., H.J., J.J.Y., H.W. and B.H.Y. provided important reagents and mouse samples.

Author Information The gene expression and ChIP data are deposited in NCBI's Gene Expression Omnibus under accession numbers GSE23743, GSE24426, GSE15179, GSE11794, GSE10086, GSE20987 and GSE24400. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.M. (markus.muschen@ucsf.edu).

METHODS

Patient samples, human cells and cell lines. Patient samples (Supplementary Table 1) were provided from the departments of Hematology and Oncology, University Hospital Benjamin Franklin, Berlin, Germany (W.-K.H.) and the USC Norris Comprehensive Cancer Center in compliance with Institutional Review Board regulations (approval from the Ethik-Kommission of the Charité, Campus Benjamin Franklin and the IRB of the University of Southern California Health Sciences Campus). Leukaemia cells from bone marrow biopsy of patients with Ph^+ ALL were xenografted into sublethally irradiated NOD/SCID mice by tail vein injection. After passaging, leukaemia cells were collected and cultured on OP9 stroma cells in alpha minimum essential medium (Alpha-MEM, Invitrogen) without ribonucleotides and deoxyribonucleotides, supplemented with 20% fetal bovine serum, 2 mmol l⁻¹ L-glutamine, 1 mmol l⁻¹ sodium pyruvate, 100 IU ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin. The human ALL cell lines BV173, NALM-1, SUP-B15 and TOM1 (obtained from Deutsche Sammlung von Microorganismen und Zellkulturen (DSMZ)) were maintained in Roswell Park Memorial Institute medium (RPMI-1640, Invitrogen) with GlutaMAX containing 20% fetal bovine serum, 100 IU ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin.

Retroviral constructs and transduction. Transfection of retroviral constructs encoding BCR-ABL1-IRES-GFP³⁰, BCR-ABL1-IRES-Neo, STAT5-CA³¹, CD44S-Puro³², FoxO4-Puro, BCL6-ER^{T2}-GFP²⁶, ER^{T2}-GFP, DN-BCL6-ER^{T2}-GFP, Cre-ER^{T2}-Puro³³, Cre-IRES-GFP, Puro-, Neo- and GFP-empty vector controls were performed using Lipofectamine 2000 (Invitrogen) with Opti-MEM media (Invitrogen). Retroviral supernatant was produced by co-transfecting HEK 293FT cells with the plasmids pHIT60 (ref. 34) (gag-pol) and pHIT123 (ecotropic env) or pHIT456 (amphotropic env). 293FT cells were cultured in high glucose Dulbecco's modified Eagle's medium (DMEM, Invitrogen) with GlutaMAX containing 10% fetal bovine serum, 100 IU ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin, 25 mmol l⁻¹ HEPES, 1 mmol l⁻¹ sodium pyruvate and 0.1 mmol l⁻¹ non-essential amino acids. Regular media were replaced after 16 h by growth media containing 10 mmol l⁻¹ sodium butyrate. After incubation for 8 h, the media were changed back to regular growth media. Twenty-four hours later, the virus supernatant was collected, filtered through a 0.45 µm filter and loaded by centrifugation (2,000g, 90 min at 32 °C) two times on 50 µg ml⁻¹ RetroNectin- (Takara) coated non-tissue six-well plates. Two million to three million cells were transduced per well by centrifugation at 500g for 30 min and maintained for 48 h at 37 °C with 5% CO₂ before transferring into culture flasks. Transduced cells with oestrogen receptor fusion proteins were induced with 4-hydroxytamoxifen (500 nM).

In vivo model for BCR-ABL1-transformed ALL and bioluminescence imaging. After cytokine-independent proliferation, BCR-ABL1-transformed ALL cells were labelled with a lentiviral vector encoding firefly luciferase with a neomycin selection marker. After selection with 0.5–2 mg ml⁻¹ G418 for 10 days, luciferase-labelled ALL cells were injected into sublethally irradiated (250 cGy) NOD/SCID mice. Human primary leukaemia cells were transduced with a lentiviral firefly luciferase carrying a GFP marker. After expansion of sorted GFP⁺ cells, 1×10^5 cells were injected through the tail vein into sublethally irradiated NOD/SCID mice. Bioimaging of leukaemia progression in mice was performed at different time points using an *in vivo* IVIS 100 bioluminescence/optical imaging system (Xenogen). D-Luciferin (Promega) dissolved in PBS was injected intraperitoneally at a dose of 2.5 mg per mouse 15 min before measuring the luminescence signal. General anaesthesia was induced with 5% isoflurane and continued during the procedure with 2% isoflurane introduced through a nose cone. All mouse experiments were subject to institutional approval by the Children's Hospital Los Angeles Institutional Animal Care and Use Committee.

Extraction of bone marrow cells from mice. To avoid inflammation-related effects in BCL6^{-/-} mice³⁵, bone marrow cells were extracted from young age-matched BCL6^{+/+} and BCL6^{-/-} mice (younger than 6 weeks of age) without signs of inflammation. Bone marrow cells were obtained by flushing cavities of femur and tibia with PBS. After filtration through a 70 µm filter and depletion of erythrocytes using a lysis buffer (BD PharmLyse, BD Biosciences), washed cells were either frozen for storage or subjected to further experiments.

BCL6^{-/-}, Stat5^{fl/fl}, Pten^{fl/fl} and p53^{-/-} mice. A summary of mouse strains used in this study is provided in Supplementary Table 3. Bone marrow cells from BCL6^{-/-} (generated in R. Dalla-Favera's laboratory)³⁶, Stat5^{fl/fl} (generated in L. Henninghausen's laboratory)³⁷, Pten^{fl/fl} (generated in H. Wu's laboratory)³⁸ and p53^{-/-} (obtained from Jackson Laboratory) mice were collected and retrovirally transformed by BCR-ABL1 (ref. 30) in the presence of 10 ng interleukin-7 per milliliter (Peprotech) in RetroNectin- (Takara) coated Petri dishes as described below. All BCR-ABL1-transformed ALL cells derived from bone marrow of mice were maintained in Iscove's modified Dulbecco's medium (IMDM, Invitrogen) with GlutaMAX containing 20% fetal bovine serum, 100 IU ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin and 50 µM 2-mercaptoethanol. BCR-ABL1-transformed ALL cells were propagated only for short periods of time and usually not longer than for

2 months to avoid acquisition of additional genetic lesions during long-term cell culture.

RI-BPI. Homo-dimerization of the amino (N)-terminal Broad Complex, Tramtrack, Bric à brac (BTB) domain of BCL6 forms a lateral groove motif, which is required to recruit co-repressor proteins such as BCL6 co-repressor (BCoR), nuclear receptor co-repressor (N-CoR) and silencing mediator of retinoid and thyroid receptors (SMRT). BCoR, NCoR and SMRT interact in a mutually exclusive manner with an 18-amino-acid motif in the lateral groove of the BCL6 BTB domain to form a BCL6 repression complex^{39,40}. A recombinant peptide containing the SMRT BBD (BCL6-binding domain) along with a cell-penetrating TAT domain was able to inhibit the transcriptional repressor activity of BCL6⁴¹. Based on this initial work, the peptidomimetic molecule RI-BPI with superior potency and stability was developed²⁵ and used for BCL6-inhibition. RI-BPI represents a retro-inverso TAT-BBD-Fu (fusogenic) peptide²⁵ that was synthesized by Biosynthesis Inc. (Lewisville, TX) and stored lyophilized at -20 °C until reconstituted with sterile, distilled, degassed water immediately before use. The purity determined by high-performance liquid chromatography-mass spectrometry was 95% or higher. RI-BPI was injected intraperitoneally into mice.

BCR-ABL1 TKI. Imatinib (STI571) and nilotinib (AMN107) were obtained from Novartis Pharmaceuticals or from LC Laboratories. Stock solutions of imatinib were prepared in sterile, distilled water at 10 mmol l⁻¹ and stored at -20 °C. Nilotinib was either dissolved in DMSO (dimethyl sulphoxide) or NMP (N-Methyl-2-pyrrolidone) just before administration. Nilotinib dissolved in DMSO was vortexed with four volumes of peanut butter until a homogeneous mixture was formed. Nilotinib (free base) solubilized in NMP was diluted with PEG 300 (polyethylene glycol 300) in a 10/90 (vol/vol) ratio. Cohorts of mice were treated with oral administration of vehicle or nilotinib (25 mg kg⁻¹ day⁻¹ or 50 mg kg⁻¹ day⁻¹) once daily at indicated time points.

Clonality analysis and spectratyping of B-cell populations. Immunoglobulin V_H-DJ_H gene rearrangements were amplified using PCR primers specific for the J558 V_H region gene with a primer specific for the C_μ constant region gene. Using a FAM-conjugated C_μ constant region or a J_H gene-specific primer in a run-off reaction, PCR products were labelled and subsequently analysed on a capillary sequencer (ABI3100, Applied Biosystems) by fragment-length analysis. Sequences of primers used are given in Supplementary Table 4.

Affymetrix GeneChip analysis. Total RNA from cells used for microarray or RT-PCR analysis was isolated by RNeasy (Qiagen) purification. RNA quality was first checked by using an Agilent Bioanalyser (Agilent Technologies). Complementary DNA (cDNA) was generated from 5 µg of total RNA using a poly(dT) oligonucleotide containing a T7 RNA polymerase initiation site and the SuperScript III Reverse Transcriptase (Invitrogen). Biotinylated cRNA was generated and fragmented according to the Affymetrix protocol and hybridized to U133A 2.0 human or 430 mouse microarrays (Affymetrix). After scanning (GeneChip Scanner 3000 7G, Affymetrix) of the GeneChip arrays, the generated CEL files were imported to BRB Array Tool (<http://linus.nci.nih.gov/BRB-ArrayTools.html>) and processed using the RMA algorithm (Robust Multi-array Average) for normalization and summarization. Relative signal intensities of probe sets were determined by comparing the signal intensity from TKI-treated and untreated cells to the average signal value of the respective cell line or a group of cell lines. The calculated signal ratios of probe sets were visualized as a heatmap with Java Treeview.

Target validation of RI-BPI in human Ph^+ ALL cells. Ph^+ ALL cell lines (BV173, NALM1 and TOM1) were treated with vehicle (control), 10 µmol l⁻¹ imatinib or imatinib + 20 µmol l⁻¹ RI-BPI for 24 h and maintained in Allprotect (Qiagen) at -80 °C until RNA isolation using an RNeasy Plus kit (Qiagen). RNA integrity was determined using the RNA 6000 Nano LabChip kit on Agilent 2100 Bioanalyser (Agilent Technologies). Two independent samples were analysed for each condition. RNA (1 µg) was hybridized to Agilent 60-mer Whole Human Genome Microarrays (part number G4112A) according to the manufacturer's recommendations. After hybridization, the processed microarrays were scanned with the Agilent DNA microarray scanner (part number G2505C) and extracted with Agilent Feature Extraction software version 8.5 (GEI-v5_10_Apr08). For computational analysis of signal, we used the dye-normalized signal after surrogate algorithm (gProcessedSignal) extracted from the .txt files and process for each array and for all the probes. This value was subjected to log₂ transformation and median array normalization. The fold changes of imatinib compared with control and (imatinib + RI-BPI) compared with imatinib were calculated for each cell line and for each gene. A data set containing previously identified BCL6 target genes (obtained from Nimblegen arrays) was mapped into the Agilent probe sets using the Agilent and NimbleGen array annotation files. To determine if two data sets differed significantly, we compared the fold change in BCL6 target genes with the fold change in BCL6 non-target genes for each data set (imatinib compared with control, and imatinib + RI-BPI compared with imatinib) by the Kolmogorov-Smirnov test⁴². The Kolmogorov-Smirnov test deter-

mines if two data sets (gene expression values for BCL6 target genes and non-target genes) differ significantly. Heat maps and other analysis were obtained using the R statistical software (<http://www.r-project.org>).

ChIP-on-chip analysis. ChIPs were performed with modifications as described⁴³. Briefly, 2.5×10^7 Ph⁺ ALL cell lines (BV173, NALM1 and TOM1) were treated with or without $10 \mu\text{mol l}^{-1}$ imatinib for 24 h. Then the cells were double cross-linked with 2 mmol l^{-1} EGS cross linker and 1% formaldehyde. After sonication, immunoprecipitations were performed using $5 \mu\text{g}$ BCL6 (N3, Santa Cruz Biotechnology) or control IgG antibody (Sigma-Aldrich) from the chromatin fragments of 2.5×10^7 human Ph⁺ ALL cells. After validation of enrichment by Q-ChIP, BCL6 or control IgG, ChIP products and their respective input genomic fragments were amplified by ligation-mediated PCR. The products were co-hybridized with the respective input samples to NimbleGen promoter arrays (human genome version 35). Quantitative ChIP was performed again at this stage for selected positive control loci to verify that the enrichment ratios were retained. The genomic products of two biological ChIP replicates were labelled with Cy5 (for ChIP products) and Cy3 (for input) and co-hybridized on custom-designed genomic tiling arrays generated by NimbleGen Systems. These high-density tiling arrays contained 50-residue oligonucleotides with an average overlap of 25 bases, omitting repetitive elements. After hybridization, the relative enrichment for each probe was calculated as the signal ratio of ChIP to input. Peaks of enrichment for BCL6 relative to input were captured with a five-probe sliding window, and the results were uploaded as custom tracks into the University of California Santa Cruz genome browser and graphically represented as histograms. Two replicates were performed with each condition.

Data analysis of ChIP-on-chip experiments. To identify target genes of BCL6 in these experiments, we computed the log-ratio between the probe intensities of the ChIP product and input and took moving averages of log-ratio of three neighbouring probes and determined the maximum value for each gene promoter and the random permutation probes as background control⁴⁴. The cut-off for each array was established as higher than the 99th percentile of the 24,175 log-ratio values generated from random permutation probes. A locus with maximum moving average above cut-offs in two replicates was considered a potential binding site. Because this high stringent-overlapping approach can produce a high false-negative rate, we also computed the correlations among peaks between the replicates to rescue promoters that did not pass cut-off in one replicate. We calculated the Pearson correlation coefficient of the probe's signal of the promoter between replicates, and promoters with a correlation higher than 0.8 were rescued and included in our final set of BCL6 targets. In addition, all peaks were mapped back to the genome using BLAT (the BLAST-like Alignment Tool, <http://genome.ucsc.edu>) to identify genes on opposite strands that could be regulated from the same bidirectional promoter. Two genes were considered to be bidirectional partners when they were located on the opposite strands in a 'head-to-head' orientation and their transcription start sites were separated by less than 0.5 kilobases.

Comparative genomic hybridization. To analyse genetic instability and acquisition of genetic lesions during long-term cell culture, genomic DNA of BCR-ABL1-transformed BCL6^{+/+} and BCL6^{-/-} ALL cells was extracted after culturing for 4 months. Genomic DNA was isolated using the PureLink genomic DNA kit (Invitrogen). Three samples of each ALL type were co-hybridized with genomic DNA extracted from normal untransformed mouse cells to NimbleGen mouse 720k Whole-Genome Tiling arrays (NimbleGen Systems) in accordance with the manufacturer's recommendations. Copy number variations were analysed using the FASST-segmentation algorithm in Nexus software (BioDiscovery). Copy-number analysis was performed using a significance threshold of 1×10^{-7} and a log₂ ratio cut-off at ± 0.2 for regions sized 1,000 kilobase pairs.

Senescence-associated β -galactosidase assay. Senescence-associated β -galactosidase activity was performed on cytospin preparations as described²¹. Briefly, a fixative solution (0.25% glutaraldehyde, 2% paraformaldehyde in PBS pH 5.5 for mouse cells and pH 6 for human cells) was freshly generated. To this end, 1 g paraformaldehyde was dissolved in 50 ml PBS at pH 5.5 by heating followed by addition of 250 μl of a 50% stock glutaraldehyde solution. $1 \times$ X-gal staining solution was prepared as follows (10 ml): 9.3 ml PBS/MgCl₂, 0.5 ml $20 \times$ KC solution (that is, 820 mg K₃Fe(CN)₆ and 1,050 mg K₄Fe(CN)₆ \times 3H₂O in 25 ml PBS) and 0.25 ml $40 \times$ X-gal (that is, 40 mg 5-bromo-4-chloro-3-indolyl β -D-galactoside per milliliter of N,N-dimethylformamide) solution were mixed. For BCR-ABL1-transformed ALL cells, 100,000 cells per cytospin were used (700 r.p.m., 8 min). The fixative solution was pipetted on the cytospins and incubated for 10 min at room temperature, then washed twice for 5 min in PBS/MgCl₂. Cytospin preparations were submerged in $1 \times$ X-gal solution, incubated overnight at 37 °C in a humidified chamber and washed twice in PBS. Slides were mounted before they dried.

Western blotting. Cells were lysed in CellLytic buffer (Sigma) supplemented with 1% protease inhibitor cocktail (Pierce). Ten micrograms of protein mixture per sample were separated on NuPAGE (Invitrogen) 4–12% Bis-Tris gradient gels and transferred on PVDF membranes (Immobilion, Millipore). To detect mouse and

human proteins by western blot, primary antibodies were used with the WesternBreeze immunodetection system (Invitrogen). The following antibodies were used: human BCL6 (clones D8 and N3, Santa Cruz Biotechnology), mouse BCL6 (rabbit polyclonal, Cell Signaling Technology), Arf (4C6/4, Cell Signaling Technology), p53 (1C12, Cell Signaling Technology), PTEN (A2B1, Santa Cruz), global Stat5 (3H7, Cell Signaling Technology) and phospho-Y694 Stat5 (14H2, Cell Signaling Technology). Antibodies against β -actin were used as a loading control (C4, Santa Cruz).

Flow cytometry. Antibodies against mouse CD19 (1D3), B220 (RA3-6B2), CD3 (17A2), CD43 (S7), CD45.1 (A20), CD45.2 (104), CD44 (IM7 and G44-26) and c-Kit (2B8) as well as respective isotype controls were purchased from BD Biosciences. For apoptosis analyses, Annexin V, propidium iodide and 7-AAD were used (BD Biosciences).

Cell viability assay. Fifty thousand BCR-ABL1-transformed ALL cells per well were seeded in a volume of 100 μl B-cell medium on Optitux 96-well plate (BD Biosciences). Imatinib was diluted in medium and added at the indicated concentration in a total culture volume of 150 μl . After culturing for 3 days, 15 μl of Resazurin (R&D) was added on each well and incubated for 4 h at 37 °C. The fluorescence was read at 535 nm and the reference wavelength was 590 nm. Fold changes were calculated using baseline values of untreated cells as a reference (set to 100%).

Colony-forming assay. The methylcellulose colony-forming assays were performed with 10,000 BCR-ABL1-transformed mouse BCL6^{-/-} or BCL6^{+/+} or 10,000 human BCR-ABL1 ALL cells. Cells were re-suspended in MethoCult medium (StemCell Technologies) and cultured on dishes (3 cm diameter) with an extra water supply dish to prevent evaporation. After 7–14 days, colonies were counted.

Cell-cycle analysis. For cell-cycle analysis in BCR-ABL1 ALL cells, the BrdU flow cytometry kit for cell-cycle analysis (BD Biosciences) was used according to manufacturer's instructions. BrdU incorporation (APC-labelled anti-BrdU antibodies) was measured with DNA content (7-amino-actinomycin-D) in fixed and permeabilized cells. The analysis was gated on viable cells that were identified based on scatter morphology⁴⁵.

In vivo toxicology studies of RI-BPI/nilotinib combinations. Fifteen adult male C57BL/6 mice were purchased from the National Cancer Institute and randomized in three groups of five. One group was exposed to intraperitoneal administration of RI-BPI 20 mg kg^{-1} of body weight three times a week; the second group was treated with RI-BPI 20 mg kg^{-1} of body weight three times a week plus nilotinib 25 mg kg^{-1} of body weight three times a week by oral gavage. A third group of five mice was treated with vehicle and used as controls. The mice were observed, examined and weighed every other day during the treatment period. Blood was collected at the end of the treatment by retro-orbital bleeding under anaesthesia. All mice were euthanized by CO₂ aspiration and the organs were harvested, weighed and macroscopically examined. Histology sections were prepared with haematoxylin and eosin staining. Pictures were taken using a digital camera (Olympus DP72) attached to a light microscope (Axioskop, Carl Zeiss) with $\times 4$ and $\times 20$ Plan Neofluar objectives (Carl Zeiss).

30. Pear, W. S. *et al.* Efficient and rapid induction of a chronic myelogenous leukemia-like myeloproliferative disease in mice receiving P210 bcr/abl-transduced bone marrow. *Blood* **92**, 3780–3792 (1998).
31. Onishi, M. *et al.* Identification and characterization of a constitutively active STAT5 mutant that promotes cell proliferation. *Mol. Cell. Biol.* **18**, 3871–3879 (1998).
32. Godar, S. *et al.* Growth-inhibitory and tumor-suppressive functions of p53 depend on its repression of CD44 expression. *Cell* **134**, 62–73 (2008).
33. Kumar, M. S. *et al.* Dicer1 functions as a haploinsufficient tumor suppressor. *Genes Dev.* **23**, 2700–2704 (2009).
34. Soneoka, Y. *et al.* A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* **23**, 628–633 (1995).
35. Dent, A. L., Shaffer, A. L., Yu, X., Allman, D. & Staudt, L. M. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science* **276**, 589–592 (1997).
36. Ye, B. H. *et al.* The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nature Genet.* **16**, 161–170 (1997).
37. Cui, Y. *et al.* Inactivation of Stat5 in mouse mammary epithelium during pregnancy reveals distinct functions in cell proliferation, survival, and differentiation. *Mol. Cell. Biol.* **24**, 8037–8047 (2004).
38. Groszer, M. *et al.* Negative regulation of neural stem/progenitor cell proliferation by the Pten tumor suppressor gene in vivo. *Science* **294**, 2186–2189 (2001).
39. Ahmad, K. F. *et al.* Mechanism of SMRT corepressor recruitment by the BCL6 BTB domain. *Mol. Cell* **12**, 1551–1564 (2003).
40. Ghetu, A. F. *et al.* Structure of a BCOR corepressor peptide in complex with the BCL6 BTB domain dimer. *Mol. Cell* **29**, 384–391 (2008).
41. Polo, J. M. *et al.* Specific peptide interference reveals BCL6 transcriptional and oncogenic mechanisms in B-cell lymphoma cells. *Nature Med.* **10**, 1329–1335 (2004).
42. Chakravarti, L. & Roy, (1967). Handbook of Methods of Applied Statistics, Volume I, John Wiley and Sons, pp. 392–394. (1967).
43. Ci, W. *et al.* The BCL6 transcriptional program features repression of multiple oncogenes in primary B cells and is deregulated in DLBCL. *Blood* **113**, 5536–5548 (2009).

44. Polo, J. M. *et al.* Transcriptional signature with differential expression of BCL6 target genes accurately identifies BCL6-dependent diffuse large B cell lymphomas. *Proc. Natl Acad. Sci. USA* **104**, 3207–3212 (2007).
45. Trageser, D. *et al.* Pre-B cell receptor-mediated cell cycle arrest in Philadelphia chromosome-positive acute lymphoblastic leukemia requires IKAROS function. *J. Exp. Med.* **206**, 1739–1753 (2009).

Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells

Hao Wu^{1†*}, Ana C. D'Alessio^{2*}, Shinsuke Ito², Kai Xia², Zhibin Wang³, Kairong Cui⁴, Keji Zhao⁴, Yi Eve Sun¹ & Yi Zhang²

Epigenetic modification of the mammalian genome by DNA methylation (5-methylcytosine) has a profound impact on chromatin structure, gene expression and maintenance of cellular identity¹. The recent demonstration that members of the Ten-eleven translocation (Tet) family of proteins can convert 5-methylcytosine to 5-hydroxymethylcytosine raised the possibility that Tet proteins are capable of establishing a distinct epigenetic state^{2,3}. We have recently demonstrated that Tet1 is specifically expressed in murine embryonic stem (ES) cells and is required for ES cell maintenance². Using chromatin immunoprecipitation coupled with high-throughput DNA sequencing, here we show in mouse ES cells that Tet1 is preferentially bound to CpG-rich sequences at promoters of both transcriptionally active and Polycomb-repressed genes. Despite an increase in levels of DNA methylation at many Tet1-binding sites, Tet1 depletion does not lead to downregulation of all the Tet1 targets. Interestingly, although Tet1-mediated promoter hypomethylation is required for maintaining the expression of a group of transcriptionally active genes, it is also involved in repression of Polycomb-targeted developmental regulators. Tet1 contributes to silencing of this group of genes by facilitating recruitment of PRC2 to CpG-rich gene promoters. Thus, our study not only establishes a role for Tet1 in modulating DNA methylation levels at CpG-rich promoters, but also reveals a dual function of Tet1 in promoting transcription of pluripotency factors as well as participating in the repression of Polycomb-targeted developmental regulators.

The Tet protein family includes three members (Tet1–3), all of which have the capacity to convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) in a 2-oxoglutarate- and Fe(II)-dependent manner^{2,3}. Consistent with the relative enrichment of 5hmC in ES cells, Tet1 is highly expressed in undifferentiated ES cells and *Tet1* messenger RNA levels decrease upon ES cell differentiation^{2,3}. Lentiviral-mediated depletion of Tet1 in mouse E14 ES cells cultured under feeder-free conditions leads to phenotypic changes that include partial loss of alkaline phosphatase activity and SSEA1 immunoreactivity, decreased self-renewal capacity and proliferation rate, downregulation of pluripotency factor *Nanog* and upregulation of differentiation genes (for example, lineage markers for trophectoderm and primitive endoderm in a subset of cells)². Thus, Tet1 may be required for mouse ES cell maintenance.

To gain insights into the mechanism by which Tet1 contributes to ES cell function, we investigated the genome-wide distribution of Tet1 in mouse ES cells by chromatin immunoprecipitation coupled with high-throughput DNA sequencing (ChIP-seq) using a highly specific Tet1 antibody (Supplementary Fig. 1a). Analysis of replicate ChIP-seq experiments identified a total of 35,564 binding sites with high confidence ($P < 10^{-8}$, or false discovery rate (FDR) of 0.01) (Supplementary Fig. 1b, c and Supplementary Table 1). In contrast, parallel experiments

using rabbit IgG did not yield specific enrichment (Fig. 1b and Supplementary Fig. 1c). Moreover, ChIP-seq analysis also indicated that Tet1 occupancy was generally reduced in fluorescence-activated cell sorting (FACS)-sorted *Tet1*-depleted ES cells (Supplementary Fig. 2a). ChIP followed by quantitative polymerase chain reaction (qPCR) analysis further confirmed decreased Tet1 occupancy on randomly selected Tet1-binding sites in response to Tet1 depletion (Supplementary Fig. 2b). Most Tet1 binding sites are located in gene-rich euchromatic regions, as 79.8% of all Tet1-bound loci are within intragenic regions or 5 kb intergenic regions up- or downstream of annotated genes (Supplementary Fig. 3a, b). Similar to other CXXC zinc-finger-domain-containing proteins (for example, Cfp1 and Kdm2a)^{4,5}, Tet1 is enriched (86.6%) at CpG islands (Fig. 1a–c). Consistently, *de novo* motif discovery analysis⁶ identified a CpG-rich sequence as the highest ranking motif within Tet1-bound regions (Fig. 1d). Quantification of CpG density within Tet1-binding loci indicated that, similar to Kdm2a (Supplementary Fig. 4a, b), Tet1 occupancy positively correlates with CpG density (Supplementary Fig. 4c). Collectively, the above results indicate that Tet1 high-affinity binding sites are generally enriched for CpG-rich sequences.

Because Tet proteins are capable of converting 5mC to 5hmC^{2,3}, we investigated the relationship between Tet1 occupancy and DNA methylation in mouse ES cells using methylated DNA immunoprecipitation coupled with mouse whole-genome tiling microarrays (MeDIP-chip). We found that DNA methylation is generally excluded from transcription start sites (TSSs) of Tet1-bound gene promoters (Fig. 2a, blue line in left panel). In contrast, Tet1-unbound gene promoters are frequently DNA methylated (Fig. 2a, red line in left panel). These results are consistent with previous studies demonstrating that CpG-rich gene promoters, where Tet1 is enriched (Fig. 1), are generally hypomethylated^{7,8}. Further analysis indicates that CpG islands not bound by Tet1 are associated with higher 5mC levels compared to Tet1-bound CpG islands (Fig. 2a, right panel). Thus, Tet1 occupancy at gene promoters is inversely correlated to levels of DNA methylation.

To investigate whether Tet1 is required for maintaining the hypomethylated state at Tet1-bound regions, we analysed DNA methylation profiles in *Tet1*-depleted ES cells and demonstrated that *Tet1* deficiency led to a general increase in 5mC levels at both TSSs and genomic regions flanking the proximal promoters of CpG-rich genes (Fig. 2b, c and Supplementary Fig. 5a, b). An increase in 5mC levels was also detected within proximal promoter regions of a subset of CpG-poor gene promoters (Fig. 2b, c and Supplementary Fig. 5a, b). The observed 5mC changes in *Tet1*-depleted cells were not due to interarray variations as a co-hybridization strategy analysing biologically independent replicates also revealed that the increase in 5mC levels induced by *Tet1* deficiency was generally enriched at Tet1-binding sites (Fig. 2d and Supplementary Table 2). Locus-specific bisulphite sequencing confirmed that Tet1-binding sites and their surrounding

¹Departments of Molecular & Medical Pharmacology and Psychiatry & Biobehavioral Sciences, IDDR at Semel Institute of Neuroscience and Human Behavior, UCLA David Geffen School of Medicine, Los Angeles, California, 90095, USA. ²Howard Hughes Medical Institute, Department of Biochemistry and Biophysics, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7295, USA. ³Laboratory of Human Environmental Epigenomes, Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA. ⁴Laboratory of Molecular Immunology, The National Heart, Lung, and Blood Institute, NIH, Bethesda, Maryland 20892, USA. [†]Present address: Cardiovascular Research Centre, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA and Department of Stem Cell and Regenerative Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA.

*These authors contributed equally to this work.

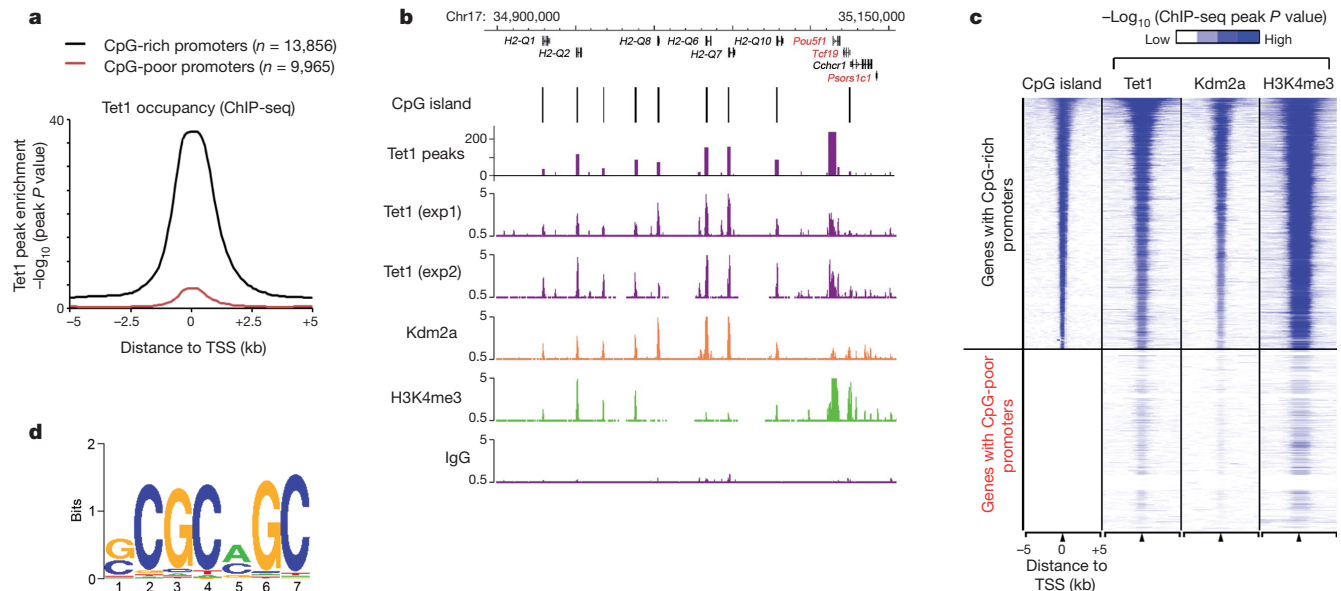


Figure 1 | Tet1 is enriched at genomic regions with high-density CpG dinucleotides. **a**, Genome-wide occupancy of Tet1 at all annotated gene promoters in ES cells (black, CpG-rich genes; red, CpG-poor genes). The enrichment of Tet1 binding was determined by ChIP-seq analysis. Average Tet1 binding measured by $-\log_{10}$ (peak P values) in 200-bp bins is shown within genomic regions covering 5 kb up- and downstream of TSSs. **b**, Enrichment of Tet1 (purple), Kdm2a (orange)⁴ and H3K4me3 (green)¹³ measured by ChIP-seq at representative genes in ES cells (black, CpG-rich; red, CpG-poor). ChIP-seq data are shown in reads per million with the y-axis floor set to 0.5 reads per million. Genomic regions with statistically significant enrichment of Tet1 binding (measured by $-\log_{10}$ (peak P values); $P < 10^{-8}$)

regions became more DNA methylated in response to Tet1 depletion (Supplementary Fig. 6). Collectively, these data suggest that Tet1 binding is required for maintaining a DNA hypomethylated state at a large cohort of CpG-rich gene promoters.

Previous studies have established a link between DNA methylation and histone methylation^{9–11}. To explore a potential relationship between Tet1 occupancy and histone modifications, we compared the binding profile of Tet1 with that of major histone modifications in mouse ES cells previously determined by ChIP-seq (Supplementary Table 3)^{12,13}. We found that histone H3 lysine 4 trimethylation (H3K4me3) is positively correlated to Tet1 binding at gene promoters,

are also indicated. **c**, Heatmap representation of genomic regions with high-density CpG sites (CpG islands), binding profiles of Tet1, Kdm2a⁴ and H3K4me3¹³ in ES cells at all annotated mouse genes promoters (5 kb flanking TSSs of Refseq genes). The heatmap is rank-ordered from genes with CpG islands of longest length to no CpG islands within 5-kb genomic regions flanking TSSs. The presence of CpG islands is shown in colour (blue, present; white, absent). ChIP-seq enrichment was measured by $-\log_{10}$ (peak P values) and is shown by colour scale. The following colour scales (white, no enrichment; blue, high enrichment) were used for Tet1/Kdm2a and H3K4me3 respectively: (0, 50) and (0, 200). **d**, A DNA motif that is enriched in Tet1-bound loci in ES cells.

as 71.3% of all Tet1-binding sites ($n = 25,359$) overlapped with H3K4me3 peaks (Fig. 1c). Analysis of the histone modification profiles that flank TSSs of Tet1-bound genes revealed two categories of Tet1 targets (Fig. 3a, b and Supplementary Table 4). The first group is associated with bivalent domains, a chromatin state characterized by the presence of both H3K4me3 and H3K27me3¹⁴. Interestingly, bivalent gene promoters in ES cells are generally hypomethylated¹⁵. In contrast, the second group is associated with active histone marks, including H3K4me3, H3K4me1 and H3K36me3 (Fig. 3a). These data indicate that Tet1 can associate with both actively transcribed as well as repressed target genes. Gene ontology analysis indicated that genes related to development and cell differentiation are highly enriched in the first group of Tet1 targets, whereas genes involved in house-keeping functions are enriched in the second group of Tet1 targets (Supplementary Fig. 7).

The fact that Tet1 occupies the promoters of actively transcribed as well as repressed genes suggests that Tet1 might have a dual function in

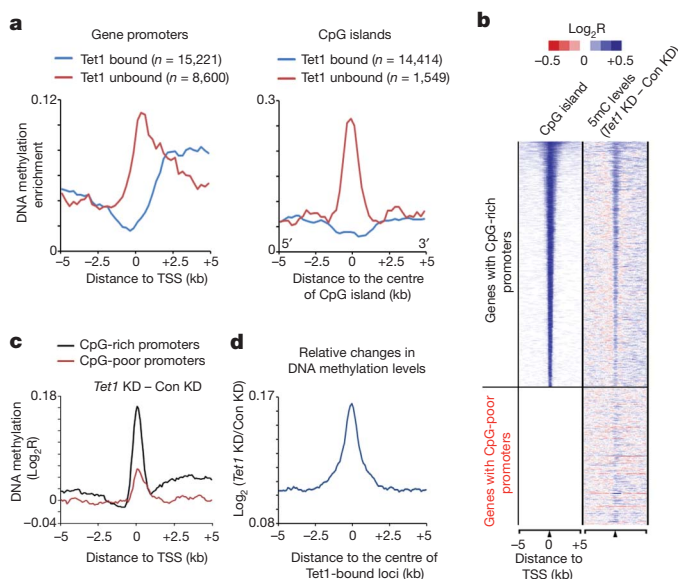


Figure 2 | Tet1 maintains a DNA hypomethylated state at Tet1-bound regions. **a**, The distribution frequency of regions enriched with DNA methylation is shown for Tet1-bound (blue) and unbound (red) gene promoters (left) or CpG islands (right) in mouse ES cells. **b**, Heatmap representation of CpG islands and the changes in DNA methylation (5mC) in response to Tet1 depletion. The DNA methylation gained after Tet1 depletion was calculated by deduction of 5mC levels in control knockdown (Con KD) from that in *Tet1* knockdown (*Tet1* KD). **c**, Changes in 5mC levels in response to *Tet1* knockdown are shown for both CpG-rich and CpG-poor gene promoters. Note that proximal promoters and 5' intragenic regions of CpG-rich genes are associated with a higher increase in 5mC levels as compared to those of CpG-poor genes in response to Tet1 depletion. **d**, An increase in 5mC levels in response to Tet1 depletion is specifically enriched at the centre of Tet1-binding loci. Changes in 5mC levels between control knockdown and *Tet1* knockdown ES cells were determined by co-hybridizing and analysing genomic DNA from control knockdown and *Tet1* knockdown cells on the whole-genome tiling microarrays.

transcription regulation. Microarray analysis comparing the gene expression of control and *Tet1*-depleted mouse ES cells identified a total of 1,332 genes that are differentially expressed (788 upregulated and 544 downregulated in *Tet1* knockdown cells) (Supplementary Fig. 8). Of these differentially expressed genes, a significant percentage (80%) are associated with Tet1 occupancy within 5 kb up- or downstream of their TSSs (1,067 out of 1,332) (Fig. 3c and Supplementary Table 5). Interestingly, despite the fact that DNA methylation has been primarily associated with transcriptional repression, more Tet1 targets are upregulated rather than downregulated in response to Tet1 depletion (677 targets are upregulated, $P = 2.0 \times 10^{-45}$, compared with 390 targets downregulated, $P = 4.1 \times 10^{-5}$, Fisher's exact test) (Fig. 3c and Supplementary Fig. 8a), indicating that Tet1 may also be involved in gene repression in mouse ES cells. Notably, genes with known functions in development and differentiation, for example, *Cdx2* (trophoblast), *Sox17* (endoderm) and *Krt8* (ectoderm), are among the upregulated Tet1 targets (Fig. 3c and Supplementary Fig. 8b). In contrast, genes related to pluripotency and ES cell functions (for example, *Nanog*, *Tcl1* and *Esrrb*) are among the downregulated Tet1 targets (Fig. 3c and Supplementary Fig. 8b). Consistent with the notion that changes in gene expression in response to Tet1 depletion are mainly due to Tet1-occupancy-mediated effects, instead of a secondary effect due to Nanog downregulation, overexpression of Nanog in *Tet1*-depleted

ES cells could only rescue a subset (~30%) of dysregulated Tet1 direct targets (Supplementary Fig. 9a, b). Notably, the rescued targets include pluripotency-related genes such as *Tcl1* and *Esrrb*. Gene expression profiling and qPCR with reverse transcription (RT-qPCR) analysis demonstrated that overexpression of Nanog rescued a subset of genes through direct (Nanog bound) or indirect (Nanog unbound) regulation (Supplementary Fig. 9a–c and Supplementary Table 6). Collectively, these results indicate that Tet1 is not only required for maintaining the expression of a subset of genes important for ES cell pluripotency, but also required for the repression of a cohort of developmental regulators.

Because many developmental regulators are repressed by Polycomb repressive complexes PRC1 and PRC2 (refs 16, 17), we sought to determine whether Tet1 might facilitate silencing of developmental regulators by promoting Polycomb repression. Comparison of our expression data sets to a published data set¹⁶ revealed that 43% of Tet1-repressed genes were also in the upregulated gene list of *Eed*-deficient ES cells, which is significantly higher than that expected by chance (43% versus 9.5%, $P = 3.59 \times 10^{-144}$, Fisher's exact test), supporting a potential role for Tet1 in PRC2-mediated repression of developmental regulators. Indeed, analysis of the histone modification states of Tet1-regulated genes in wild-type ES cells indicated that Tet1-repressed genes were preferentially associated with H3K27me3 (Fig. 3c), a mark deposited by

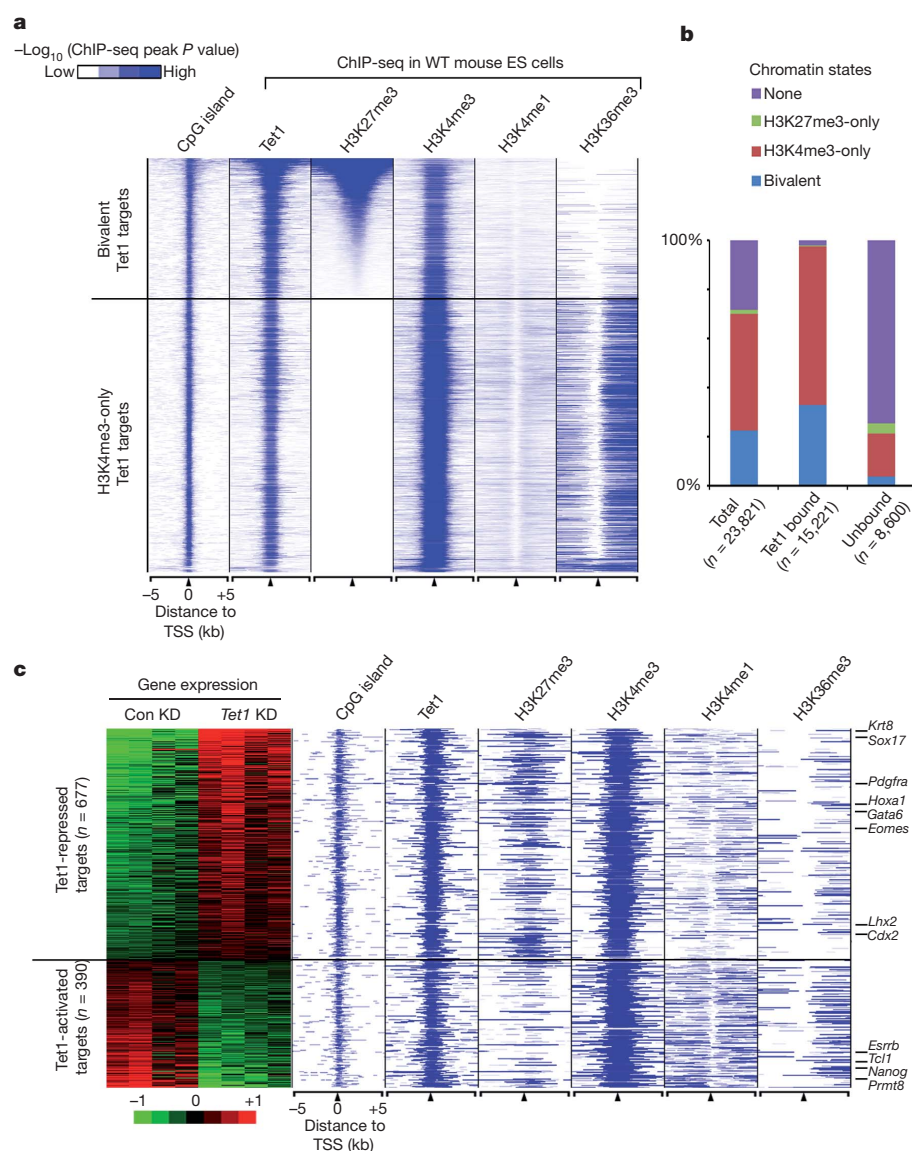


Figure 3 | Tet1 binds to and functions in both repressed (bivalent) and actively transcribed (H3K4me3-only) genes. **a**, Heatmap representation of genomic regions with high-density CpG sites (CpG islands), binding profile of Tet1, and major histone modifications (H3K4me1 (ref. 12), H3K4me3, H3K27me3 and H3K36me3 (ref. 13)) in mouse ES cells at indicated Tet1 target genes (5 kb flanking TSSs). The heatmap is rank-ordered from genes with highest H3K27me3 enrichment to no H3K27me3 within 5-kb genomic regions flanking TSSs. The following colour scales (white, no enrichment; blue, high enrichment) were used for Tet1/H3K27me3/H3K4me1/H3K36me3 and H3K4me3 respectively: (0, 50) and (0, 200). WT, wild type. **b**, Relative percentage of genes with different chromatin states shown for all genes, Tet1-bound and unbound genes. **c**, Heatmap representation of differentially expressed Tet1 targets between control knockdown and *Tet1* knockdown mouse ES cells. Note that Tet1-repressed targets are preferentially associated with bivalent chromatin states, whereas Tet1-activated targets are generally H3K4me3-only genes.

PRC2^{18,19}. In contrast, Tet1-activated targets were preferentially associated with H3K36me3, a mark associated with transcriptional elongation²⁰ (Fig. 3c), supporting the notion that Tet1-mediated DNA hypomethylation at these gene promoters may facilitate their expression.

The fact that genes upregulated in response to *Tet1* knockdown significantly overlap with those upregulated by *Eed* deficiency indicates that Tet1 may cooperate with PRC2 in silencing this group of genes. Given that the protein levels of PRC2 subunits are not significantly altered in response to Tet1 depletion (Supplementary Fig. 10), Tet1 is unlikely to affect PRC2 expression or stability. As 95.2% of PRC2-binding sites (defined as Ezh2/Suz12 co-bound²¹) overlapped with Tet1-bound loci (Fig. 4a), we next evaluated the effect of Tet1 depletion on the chromatin-binding ability of PRC2. ChIP coupled with whole genome tiling microarrays (ChIP-chip) in control and *Tet1* knockdown cells revealed that Tet1 depletion impaired the binding of Ezh2, a core subunit of PRC2, to a large fraction (72.2%) of PRC2-binding sites (Fig. 4a, b, Supplementary Figs 11, 12a and Supplementary Table 7). ChIP-qPCR further confirmed the effect of *Tet1* knockdown on Ezh2/Suz12 recruitment (Fig. 4c and Supplementary Fig. 12b). Interestingly, depletion of Ezh2 did not affect Tet1 binding to chromatin (Fig. 4c), indicating that Tet1 may function upstream of PRC2. Furthermore,

overexpression of Nanog in *Tet1*-depleted cells also failed to fully rescue the Ezh2 binding to Tet1/PRC2 co-bound targets (Supplementary Fig. 9d). Given that previous purification of the PRC2 complex did not uncover Tet1 as an associated component^{18,19,22} and the fact that a stable interaction between Tet1 and PRC2 could not be demonstrated (unpublished observation), we favour a model in which Tet1 may indirectly contribute to PRC2 recruitment by maintaining a DNA hypomethylated state at PRC2-bound loci. This model is supported by a recent study demonstrating that DNA methylation impedes binding of PRC2 to chromatin²³.

In summary, we demonstrate that Tet1 is preferentially enriched in CpG-island-containing gene promoters in mouse ES cells. This result is consistent with the presence of a CXXC domain in Tet1 and the demonstration that the CXXC domain is preferentially bound to CpG-rich sequences^{4,5}. The nonrandom genomic distribution of Tet1 suggests that genes with CpG-rich promoters are selectively regulated by a Tet1-dependent epigenetic state (that is, 5hmC) or active demethylation process. The convergence of CpG-binding proteins at CpG islands, including Cfp1, Kdm2a and Tet1, cooperatively contributes to the establishment of a specialized chromatin/epigenetic state at CpG-rich gene promoters. Specifically, Cfp1 confers H3K4me3 by

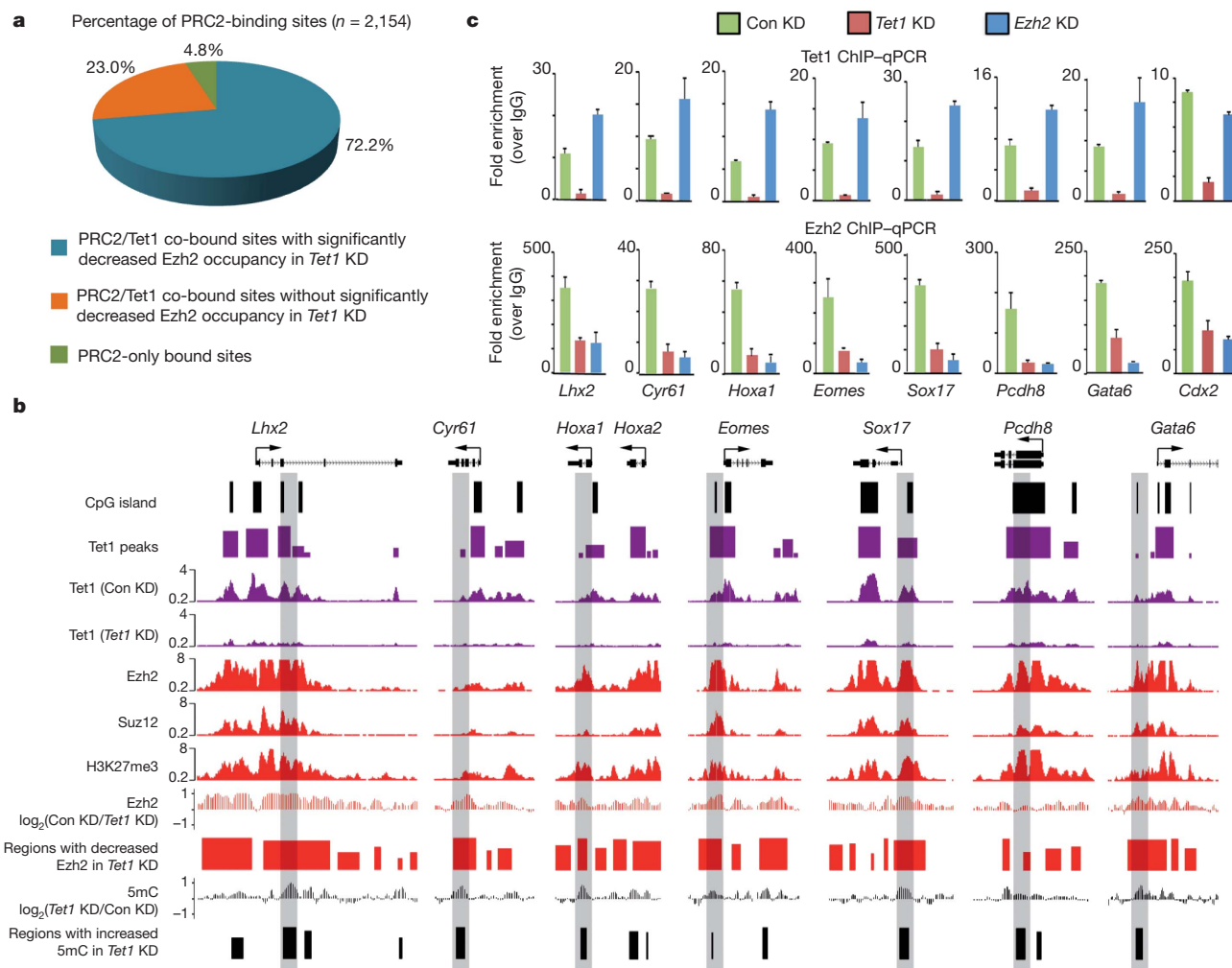


Figure 4 | Tet1 is required for chromatin binding of PRC2 in mouse ES cells. **a**, Tet1 depletion affects the binding of PRC2 to the majority of its targets. PRC2-binding sites are divided into three groups (Tet1/PRC2 co-bound Tet1 dependent, Tet1 independent and PRC2-only bound). **b**, Shown are Tet1, Ezh2 and Suz12 (ref. 21), and H3K27me3 (ref. 13) occupancy, and the effect of Tet1 depletion on Ezh2 occupancy and 5mC levels at seven representative Tet1-repressed bivalent targets. Regions associated with significant changes in Ezh2

occupancy between control and *Tet1*-depleted ES cells were measured by whole-genome tiling microarrays. Genomic regions that are further examined by locus-specific ChIP-qPCR in **c** are shaded. **c**, ChIP-qPCR analysis of Tet1 (top panels) and Ezh2 (bottom panels) occupancy at the promoters of eight representative Tet1-repressed targets in control (Con KD), *Tet1*-depleted (*Tet1* KD) and *Ezh2*-depleted (*Ezh2* KD) ES cells. Error bars represent standard deviation determined from duplicate experiments.

recruiting the H3K4me3 methyltransferase Setd1 (ref. 5); Kdm2a leads to depletion of H3K36me2 (ref. 4), and Tet1 maintains DNA at a hypomethylation state at CpG islands (Fig. 2).

In addition to binding to gene promoters with CpG islands, Tet1 also binds to a subset of actively transcribed CpG-poor gene promoters, such as *Nanog*, *Tcl1* and *Esrrb*, whose gene products have an important role in ES cell maintenance. In this scenario, Tet1 has an important role in promoting the transcriptionally active state of these genes by maintaining a hypomethylated promoter state². Interestingly, Tet1 also contributes to the silencing of a group of developmental regulators and somatic lineage differentiation genes that are silenced by Polycomb group proteins (Fig. 3c). Depletion of Tet1 leads to a decrease in *Ezh2* occupancy at many PRC2 targets, indicating that Tet1 contributes to PRC2 recruitment. Therefore, our study reveals a novel function for Tet1 in the recruitment of PRC2 and silencing of developmental regulators, which also contributes to the role of Tet1 in mouse ES cell maintenance. We note that, in contrast to our results, a recent study has shown that knockdown of *Tet1* alone is not sufficient to confer any noticeable phenotype in mouse ES cells²⁴. This difference is probably due to the use of different ES cell lines, culture conditions and knockdown efficiency (see Supplementary Information for details). Collectively, our study establishes a dual function for Tet1 in transcriptional regulation in mouse ES cells.

METHODS SUMMARY

Mouse ES cell cultures and lentiviral knockdown. Mouse E14Tg2A ES cells were cultured in feeder-free conditions². For *Tet1* knockdown, mouse ES cells were infected with lentiviruses expressing both the GFP reporter and short-hairpin RNA (shRNA) specific for *Tet1* (5'-GCAGATGCCGTGACACAAAT-3'). For *Ezh2* knockdown, mouse ES cells were infected with lentiviruses expressing both the GFP reporter and shRNA specific for *Ezh2* (5'-GTATGTGGGCATCGAACGA-3') as previously described²⁵. All analyses were performed using *Tet1*- or *Ezh2*-depleted ES cells that were purified on the basis of GFP fluorescence by FACS 8 days after lentiviral transduction. Lentiviruses expressing GFP alone was used as a control.

ChIP-seq and data analysis. ChIP and sequencing experiments were performed as described²⁶. Briefly, cells were cross-linked with 1% formaldehyde at 25 °C for 10 min and sonicated to generate chromatin fragments of 200–500 bp. Chromatin fragments from 10–20 million cells were immunoprecipitated using 8 µg of the Tet1 antibody² or IgG control from two biologically independent samples. ChIP-seq library construction and Illumina sequencing were performed as described previously²⁶. All sequencing reads were mapped to the mouse genome (NCBI Build 36/UCSC mm8). Sequencing reads from two independent Tet1 ChIP-seq experiments were combined and Tet1-enriched regions were identified by the MACS program²⁷. Sequencing reads from IgG control experiments were used as negative controls in MACS. The statistical cutoff used for identifying Tet1-binding sites was a *P* value < 10⁻⁸ and fold enrichment (over IgG control) > 10.

Genome-wide DNA methylation (5mC) analysis. Methylated DNA immunoprecipitation (MeDIP) coupled with whole-genome DNA tiling microarrays were performed as described²³. Immunoprecipitated DNA was prepared from both control and *Tet1*-depleted ES cells, and hybridized to mouse whole-genome tiling microarrays (NimbleGen).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 July 2010; accepted 14 February 2011.

Published online 30 March 2011.

1. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nature Rev. Genet.* **2008**, 129–140 (2008).
2. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
3. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).

4. Blackledge, N. P. *et al.* CpG islands recruit a histone H3 lysine 36 demethylase. *Mol. Cell* **38**, 179–190 (2010).
5. Thomson, J. P. *et al.* CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**, 1082–1086 (2010).
6. Liu, X. S., Brutlag, D. L. & Liu, J. S. An algorithm for finding protein–DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. *Nature Biotechnol.* **20**, 835–839 (2002).
7. Fouse, S. D. *et al.* Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2**, 160–169 (2008).
8. Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genet.* **39**, 457–466 (2007).
9. Mohn, F. *et al.* Lineage-specific polycomb targets and *de novo* DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30**, 755–766 (2008).
10. Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* **448**, 714–717 (2007).
11. Schlesinger, Y. *et al.* Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer. *Nature Genet.* **39**, 232–236 (2007).
12. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
13. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
14. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
15. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
16. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
17. Lee, T. I. *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301–313 (2006).
18. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
19. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16**, 2893–2905 (2002).
20. Xiao, T. *et al.* Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev.* **17**, 654–663 (2003).
21. Ku, M. *et al.* Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* **4**, e1000242 (2008).
22. Pasini, D. *et al.* JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature* **464**, 306–310 (2010).
23. Wu, H. *et al.* Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* **329**, 444–448 (2010).
24. Koh, K. P. *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
25. Wu, S. C., Kallin, E. M. & Zhang, Y. Role of H3K27 methylation in the regulation of lncRNA expression. *Cell Res.* **20**, 1109–1116 (2010).
26. Wang, Z. *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019–1031 (2009).
27. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Abraham and I. Chepelev for Illumina sequencing and data transfer; J. He and A. T. Nguyen for FACS sorting; O. Taranova for discussion; S. Wu for critical reading of the manuscript. This work was supported by NIH grants GM68804 (to Y.Z.), R56MH082068 (to Y.E.S.) and support from the Division of Intramural Research Program of National Heart, Lung and Blood Institute, NIH (K.Z.). S.I. is a research fellow of the Japan Society for the Promotion of Science. Y.Z. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions Y.Z. conceived the project; H.W., A.C.D.A. and Y.Z. designed the experiments; H.W., A.C.D.A., S.I., Z.W. and K.C. performed the experiments; H.W. and K.X. analysed the data; H.W., A.C.D.A., K.Z., Y.E.S. and Y.Z. interpreted the data; H.W. and Y.Z. wrote the manuscript.

Author Information ChIP-seq and microarray data have been deposited in the Gene Expression Omnibus under accession number GSE26833. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.Z. (yi_zhang@med.unc.edu).

METHODS

Constructs and antibodies. All the constructs and antibodies used in this study have been described previously^{2,18} or were purchased from the following sources: EZH2 (Cell Signaling; catalogue no. 4905); EED (Santa Cruz; sc-133537); Jarid2 (Abcam; ab48137); AEBP2 (Proteintech group; 11232-2-AP), EZH1 (Abcam; ab64850) and actin (Sigma; AC-40).

Mouse ES cell cultures and lentiviral knockdown. Mouse E14Tg2A ES cells were cultured in feeder-free conditions². For *Tet1* knockdown, mouse ES cells were infected with lentiviruses expressing both the GFP reporter and short-hairpin RNA (shRNA) specific for *Tet1* (5'-GCAGATGGCCGTGACACAAAT-3'). For *Ezh2* knockdown, mouse ES cells were infected with lentiviruses expressing both the GFP reporter and shRNA specific for *Ezh2* (5'-GTATGTGGGCATCGAACGA-3') as previously described²⁵. All analyses were performed using *Tet1*- or *Ezh2*-depleted ES cells that were purified on the basis of GFP fluorescence by FACS 8 days after lentiviral transduction. Lentiviruses expressing GFP alone was used as a control.

RNA isolation, qPCR and expression microarray analysis. Total RNA from cultured cells was isolated using RNeasy Mini Kit (Qiagen), and cDNA was generated with Improm-ITM Reverse Transcription System (Promega). Real-time qPCR reactions were performed on an ABI PRISM 7700 Sequence Detection System (Applied Biosystems) using SYBR Green (Invitrogen). cDNA levels of target genes were analysed using comparative C_T methods, where C_T is the cycle threshold number and normalized to GAPDH. RT-qPCR primers are listed in Supplementary Table 8.

For expression microarray analysis comparing control and *Tet1* knockdown ES cells, 2 µg of total RNA purified from GFP sorted cells were reverse-transcribed into cDNA with a T7-(dT)24 primer from a custom kit (Life Technologies). Biotinylated cRNA was then generated from the cDNA reaction using the BioArray High Yield RNA Transcript Kit. The cRNA was then fragmented in fragmentation buffer (40 mM Tris-acetate, pH 8.1, 100 mM KOAc and 150 mM MgOAc) at 94 °C for 35 min before microarray hybridization. Fifteen micrograms of fragmented cRNA was then added to a hybridization cocktail (0.05 mg ml⁻¹ fragmented cRNA, 50 pM control oligonucleotide B2, BioB, BioC, BioD and cre hybridization controls, 0.1 mg ml⁻¹ herring sperm DNA, 0.5 mg ml⁻¹ acetylated BSA, 100 mM MES, 1 M Na⁺, 20 mM EDTA, 0.01% Tween 20). Ten micrograms of cRNA were used for hybridization to Affymetrix GeneChip Mouse Genome 430 2.0 Array. Hybridization was carried out at 45 °C for 16 h. The arrays were then washed and stained with R-phycoerythrin streptavidin, before scanning. Washing, scanning and basic analysis was carried out using Affymetrix GeneChip Microarray Suite 5.0 software. Raw signal intensity (.cel files) was RMA normalized using affy (R/bioconductor). For identification of differentially expressed genes, we used NIA array analysis tool (<http://lgsun.grc.nia.nih.gov/ANOVA>). Of all the probes present on the microarray, signal intensity of redundant probes was averaged before analysis. The following parameters were used for analysing statistically significant differential expression: threshold z -value to remove outliers, 10,000; Error Model, Max (Average, Bayesian); error variance averaging window, 200; proportion of highest error variances to be removed, 0.05; Bayesian degrees of freedom, 20; the FDR threshold was set at 0.05.

For heatmap display, RMA-normalized signal intensity was log₂ transformed and median-centred. Heatmaps were generated using Cluster3 and Java Treeview. **ChIP-seq.** ChIP-seq experiments were performed as described²⁶. Briefly, cells were cross-linked with 1% formaldehyde at 25 °C for 10 min and sonicated to generate chromatin fragments of 200–500 bp. Chromatin fragments from 10–20 × 10⁶ cells were immunoprecipitated using 8 µg of the Tet1 antibody² or IgG control from two biologically independent samples. ChIP-seq library construction and Illumina sequencing were performed as described previously²⁶. All sequencing reads were mapped to the mouse genome (mm8). Sequencing reads from both Tet1 ChIP-seq experiments were combined and Tet1-enriched regions were determined by the MACS program (version 1.3.7.1). Sequencing reads from IgG control experiments were used as negative controls in MACS. Only uniquely mapped reads were retained and redundant reads were filtered out. The statistical cutoff used for identifying Tet1-binding sites was P value < 10⁻⁸ (or FDR < 1%) and fold enrichment (over IgG control) > 10. ChIP-seq data sets of H3K4me1 (ref. 12), H3K4me3, H3K27me3, H3K36me3 (ref. 13), Ezh2, Suz12 (ref. 21), Kdm2a (ref. 4) and RNA pol II (ref. 28) were obtained from previous publications and reanalysed in MACS using identical parameters (except statistical cutoff was set to P value < 10⁻⁵). A summary of all ChIP-seq experiments used in this study (generated by this work and by previous publications) is provided in Supplementary Table 3. ChIP-seq sequencing read counts for each ChIP-seq experiment were binned into 400-bp windows at 100-bp steps along the genome and visualized in the Cisgenome browser²⁹. To assign ChIP-seq enriched regions to genes, a complete set of Refseq genes was downloaded from the UCSC table browser (accessed May, 2010). For all data sets, genes with enriched regions within 5 kb of their TSSs were called bound.

Gene ontology analysis. Functional enrichment analysis of bivalent and H3K4me3-only Tet1 were calculated by hypergeometric distribution followed by Benjamini correction in DAVID.

Genome-wide DNA methylation (5mC) analysis. Methylated DNA immunoprecipitation (MeDIP) was performed as described previously with minor modifications⁸. Briefly, genomic DNA was sequentially digested with proteinase K and RNase A, and purified by phenol/chloroform extraction. Purified genomic DNA was sonicated and heat denatured (95 °C, 10 min). An aliquot of sonicated genomic DNA was saved as input. Five micrograms of fragmented genomic DNA was immunoprecipitated with 5 µl of a monoclonal antibody against 5-methylcytidine (Eurogentec) at 4 °C overnight in a final volume of 500 µl of IP buffer (10 mM sodium phosphate (pH 7.0), 140 mM NaCl, 0.05% Triton X-100). We incubated the DNA-antibody mixture with 30 µl protein G Dynabeads (Invitrogen) for 2 h at 4 °C and washed it three times with 1 ml IP buffer. We then treated the beads with proteinase K for at least 3 h at 55 °C and purified the methylated DNA by phenol-chloroform extraction followed by ethanol precipitation. For whole-genome DNA tiling microarray analysis, immunoprecipitated DNA prepared from both control and *Tet1*-depleted ES cells were co-hybridized to mouse whole-genome tiling microarrays (NimbleGen).

Whole-genome tiling microarray analysis. For whole-genome DNA tiling microarray analysis of relative changes in 5mC levels or Ezh2 occupancy, immunoprecipitated DNA was prepared from both control and *Tet1*-depleted ES cells and amplified using whole genome amplification kit (Sigma). Amplified DNA was labelled (5' Cy5- or Cy3-random nonamers, TriLink Biotechnologies) using the standard protocol (NimbleGen Arrays User's Guide for ChIP-chip analysis). Hybridization of labelled samples to whole genome HD2 microarrays 4-array set (Roche/NimbleGen, ~2.1 million tiling probes per array, covering the entire non-repetitive portion of mouse genome) was carried out for 16–20 h at 42 °C using NimbleGen hybridization System 4. After stringent washes, microarrays were subsequently scanned using an Agilent scanner at 5-µm resolution. Data were extracted and analysed using NimbleScan v2.5 (Roche/NimbleGen).

For identification of probes associated with significant increase in 5mC levels or decrease in Ezh2 occupancy in response to Tet1 depletion in microarray experiments with the IP/IP configuration (DNA from control knockdown and *Tet1* knockdown were co-hybridized to the same microarrays), a non-parametric one-sided Kolmogorov-Smirnov (KS) test was used (KS score). Briefly, from the scaled log₂-ratio data, a fixed-length window (750 bp) is placed around each consecutive probe and the one-sided KS test is applied to determine whether the probes are drawn from a significantly more positive distribution of intensity log-ratios than those in the rest of the array. The resulting score for each probe is the $-\log_{10} P$ value from the windowed KS test around that probe. Using NimbleScan v2.5, peak data files are generated from the P -value data files. NimbleScan software detects peaks by searching for at least 2 probes above a P -value minimum cutoff ($-\log_{10}$) of 2. Peaks within 500 bp of each other are merged. For calculating the absolute 5mC levels in control knockdown and *Tet1* knockdown ES cells (Supplementary Fig. 5a), the MEDME program³⁰ was used to correct the nonlinear relationship between MeDIP-chip signals (measured by microarray experiments with the IP/input configuration) and genomic CpG density.

For visualizing raw microarray signal intensity in the genome browser, probe level smoothing (log₂ ratios of probes within 1 kb are averaged) was performed for each probe. For calculating the peak distribution, regions associated with significant changes in 5mC levels or Ezh2 occupancy were binned to 500-bp intervals using a 250-bp sliding window within genomic regions 5-kb up- and downstream of TSSs of annotated Refseq genes. Heatmaps were generated and visualized using Cluster3 and Java TreeView, respectively.

Locus-specific ChIP assays and bisulphite sequencing. Cells were fixed in a final concentration of 1% formaldehyde. After incubation at 25 °C for 10 min, the reaction was stopped by the addition of 125 mM glycine. ChIP assays were performed using a protocol associated with the ChIP assay kit (Upstate Biotechnology). After extensive washing, ChIPed DNA was eluted from the beads, and analysed on an ABI 7300 Real Time PCR System (Applied Biosystems) using SYBR Green (Invitrogen). Primer sequences are listed in Supplementary Table 9.

Bisulphite sequencing was performed as described previously with minor modifications². Five micrograms of sodium-bisulphite-treated DNA samples was subjected to PCR amplification using the first set of primers; PCR products were used as templates for a subsequent PCR reaction using nested primers. The PCR products of the second reaction were then subcloned using the Invitrogen TA cloning Kit following the manufacturer's instructions. PCRs and subcloning were performed in duplicate for each sample. The clones were sequenced using the M13 reverse primer. Primers for bisulphite sequencing are listed in Supplementary Table 10.

28. Seila, A. C. *et al.* Divergent transcription from active promoters. *Science* **322**, 1849–1851 (2008).

29. Ji, H. *et al.* An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nature Biotechnol.* **26**, 1293–1300 (2008).

30. Pelizzola, M. *et al.* MEDME: an experimental and analytical methodology for the estimation of DNA methylation levels based on microarray derived MeDIP-enrichment. *Genome Res.* **18**, 1652–1659 (2008).

Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells

William A. Pastor^{1,2*}, Utz J. Pape^{1,3*}, Yun Huang^{2*}, Hope R. Henderson^{1,2}, Ryan Lister⁴, Myunggon Ko², Erin M. McLoughlin⁵, Yevgeny Brudno⁶, Sahasransu Mahapatra², Philipp Kapranov⁷, Mamta Tahiliani^{1†}, George Q. Daley⁵, X. Shirley Liu³, Joseph R. Ecker⁴, Patrice M. Milos⁷, Suneet Agarwal⁵ & Anjana Rao^{1,2}

5-hydroxymethylcytosine (5hmC) is a modified base present at low levels in diverse cell types in mammals^{1–5}. 5hmC is generated by the TET family of Fe(II) and 2-oxoglutarate-dependent enzymes through oxidation of 5-methylcytosine (5mC)^{1,2,4–7}. 5hmC and TET proteins have been implicated in stem cell biology and cancer^{1,4,5,8,9}, but information on the genome-wide distribution of 5hmC is limited. Here we describe two novel and specific approaches to profile the genomic localization of 5hmC. The first approach, termed GLIB (glucosylation, periodate oxidation, biotinylation) uses a combination of enzymatic and chemical steps to isolate DNA fragments containing as few as a single 5hmC. The second approach involves conversion of 5hmC to cytosine 5-methylenesulphonate (CMS) by treatment of genomic DNA with sodium bisulphite, followed by immunoprecipitation of CMS-containing DNA with a specific anti-serum to CMS⁵. High-throughput sequencing of 5hmC-containing DNA from mouse embryonic stem (ES) cells showed strong enrichment within exons and near transcriptional start sites. 5hmC was especially enriched at the start sites of genes whose promoters bear dual histone 3 lysine 27 trimethylation (H3K27me3) and histone 3 lysine 4 trimethylation (H3K4me3) marks. Our results indicate that 5hmC has a probable role in transcriptional regulation, and suggest a model in which 5hmC contributes to the ‘poised’ chromatin signature found at developmentally-regulated genes in ES cells.

We developed two independent methods for precipitation of 5hmC in genomic DNA. The GLIB method (Fig. 1a) entails addition of a glucose molecule to each 5hmC with T4 phage β -glucosyltransferase (BGT)³ (Supplementary Fig. 1a). The glucose moiety is oxidized with sodium periodate, which converts the vicinal hydroxyl groups to aldehydes¹⁰, and further modified with aldehyde-reactive probe, which adds two biotin molecules to each 5hmC (Fig. 1a). A related strategy, which uses a custom-synthesized UDP-glucose analogue (UDP-6-N₃-glucose), was recently used to profile 5hmC distribution in mouse brain¹¹. The second method uses an antibody against cytosine 5-methylenesulphonate (CMS)⁵, produced by reaction of 5hmC with sodium bisulphite (Fig. 1b)¹². Anti-CMS antibodies are more sensitive and less density-dependent than anti-5hmC in DNA dot blot assays⁵. Both methods are specific for DNA containing 5hmC (Supplementary Fig. 1b)⁵.

We examined the ability of GLIB-treated (biotinylated) and bisulphite-treated 5hmC-containing DNA to be pulled down by streptavidin and anti-CMS antisera, respectively. Using varying ratios of dCTP:dhmCTP, we generated 201 base pairs PCR amplicons with differing incorporation of cytosine and 5hmC in identical sequence contexts (Supplementary Table 1). At each dhmCTP:dCTP ratio, the fraction of amplicons that contain no 5hmC, and therefore should not be precipitated, can be calculated using the binomial equation (Supplementary Table 2). Observed and calculated pull-down efficiencies

were very similar (Fig. 1c): even at low densities of 5hmC, more than 90% of DNA fragments calculated to contain a single 5hmC were precipitated after GLIB treatment. Anti-CMS pull-down showed increased density dependence compared to GLIB, but had very low background, such that there was still a strong preference for precipitation of sparsely hydroxymethylated amplicons over unmodified ones (Fig. 1d). The performance of a commercial polyclonal anti-hmC antiserum was inferior to that of anti-CMS, in terms of higher background pull-down of unmodified DNA (3.0% versus 0.06%) as well as greater density dependence (Fig. 1e). By testing PCR amplicons with varying 5mC, we confirmed that the methyl-DNA immunoprecipitation (MeDIP) technique, which uses a monoclonal antibody to 5mC, is extremely density-dependent (Fig. 1f).

We applied the GLIB and anti-CMS techniques to enrich 5hmC-containing regions in genomic DNA using genomic DNA with low, intermediate and high levels of 5hmC (Supplementary Fig. 1c, left panel). For the GLIB and anti-CMS pull-downs, the amount of specifically precipitated genomic DNA was proportional to the relative amount of 5hmC (Supplementary Fig. 1c). The GLIB technique did not produce mutations (Supplementary Table 3), the biotinylated DNA could be efficiently eluted by heating with formamide (Supplementary Fig. 1d), and the biotinylated adduct had a minimal inhibitory effect on PCR at 5–25% hmC density (delay of approximately 0.1 cycles per converted 5hmC residue; Supplementary Fig. 1f). There was no PCR delay with CMS-containing PCR amplicons except at very high CMS levels (Supplementary Fig. 1e), consistent with our previous report that CMS inhibits PCR predominantly at biologically irrelevant sequences where multiple CMS adducts occur in a row¹³.

We investigated the genome-wide localization of 5hmC in murine V6.5 ES cells. For GLIB-treated DNA, we chose Helicos single molecule DNA sequencing, which does not require an amplification step and thus avoids PCR bias^{14,15}. For CMS-enriched genomic DNA, we used an Illumina instrument, as longer read lengths are needed for efficient alignment of bisulphite-treated DNA to the genome¹⁶. With the GLIB method, 119,600 regions of the genome, averaging 1,422 bp in length, showed a substantially higher density of reads in the +BGT as opposed to the –BGT sample; with the CMS method, comparison of enriched to input DNA identified 109,264 enriched regions (average length 1,168 bp). There was high overlap in the enriched regions, here designated 5hmC-enriched regions of the genome (HERGs) (Fig. 1g). Comparing the number of HERGs retrieved by using different fractions of aligned reads yielded a curve that approached an asymptote, suggesting that a majority of hydroxymethylated regions had been identified (Supplementary Fig. 2a).

To determine whether HERGs overlapped with methylated DNA regions, we identified 62,991 5mC-enriched regions of the genome

¹Harvard Medical School, Immune Disease Institute and Program in Cellular and Molecular Medicine, Children’s Hospital Boston, Boston, Massachusetts 02115, USA. ²La Jolla Institute for Allergy & Immunology, La Jolla, California 92037, USA. ³Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, Massachusetts 02115, USA. ⁴Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA. ⁵Division of Hematology/Oncology, Children’s Hospital Boston; Dana-Farber Cancer Institute; Harvard Stem Cell Institute, Boston, Massachusetts 02115, USA. ⁶Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA. ⁷Helicos BioSciences Corporation, Cambridge, Massachusetts 02139, USA. [†]Present address: Department of Biochemistry, New York University Langone Medical Centre, New York, New York 10016, USA.

*These authors contributed equally to this work.

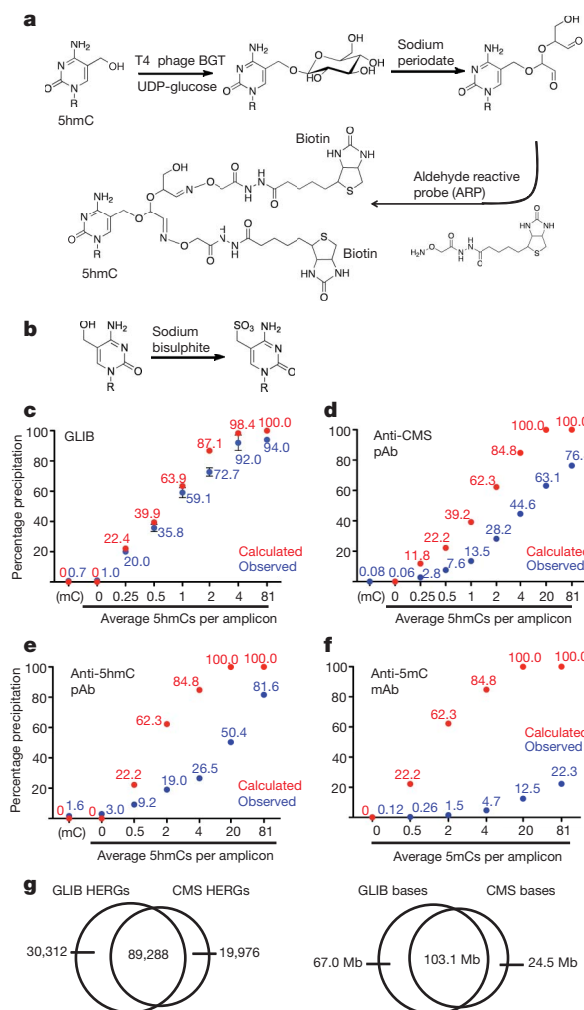


Figure 1 | Comparison of 5hmC enrichment methods. **a**, The GLIB method. Glucose is added to 5hmC by BGT, oxidized with sodium periodate to yield aldehydes, and reacted with the aldehyde reactive probe (ARP), yielding two biotins at the site of every 5hmC. **b**, 5hmC is converted to CMS by sodium bisulphite. **c–f**, Precipitation of PCR amplicons containing (1) varying amounts of 5hmC by GLIB methodology (**c**), anti-CMS methodology (**d**), or anti-5hmC antibody (**e**); or (2) varying amounts of 5mC by anti-5mC antibody (**f**). pAb, polyclonal antibody; mAb, monoclonal antibody. Between 1 and 6 independent experiments per method, mean percentage input precipitated \pm s.d. is indicated. **g**, Overlap between HERGs identified by the GLIB and anti-CMS methodologies. Left panel, number of HERGs; right panel, number of base pairs contained within HERGs.

(MERGs) by MeDIP. The resulting 5mC profile does not represent a complete map of 5mC in mouse ES cells, but rather is biased towards regions of dense methylation. Statistics pertaining to the GLIB, anti-CMS and MeDIP enrichments are shown in Supplementary Figs 2b–d, the corresponding annotations are provided in Supplementary Tables 4–9, and reads and enrichment for the *Hoxb* locus are provided in Supplementary Table 10. As expected, both HERGs and MERGs contained a high frequency of CG sequences relative to the genome at large (Supplementary Fig. 3a). Intriguingly, HERGs also contained relatively high levels of CAG sequences, the most frequent site of non-CpG methylation in human ES cells¹⁶, and we confirmed that the TET1 catalytic domain is capable of hydroxylating 5mC in CHG and CHH (H = A, T or C) contexts *in vitro* (Supplementary Fig. 3b).

Analysis of the GLIB and anti-CMS HERG sets gave very similar results. We observed a strong correlation between the densities of HERGs and genes on a given chromosome; this trend was less pronounced for MERGs (Fig. 2a). When we compared the distribution of

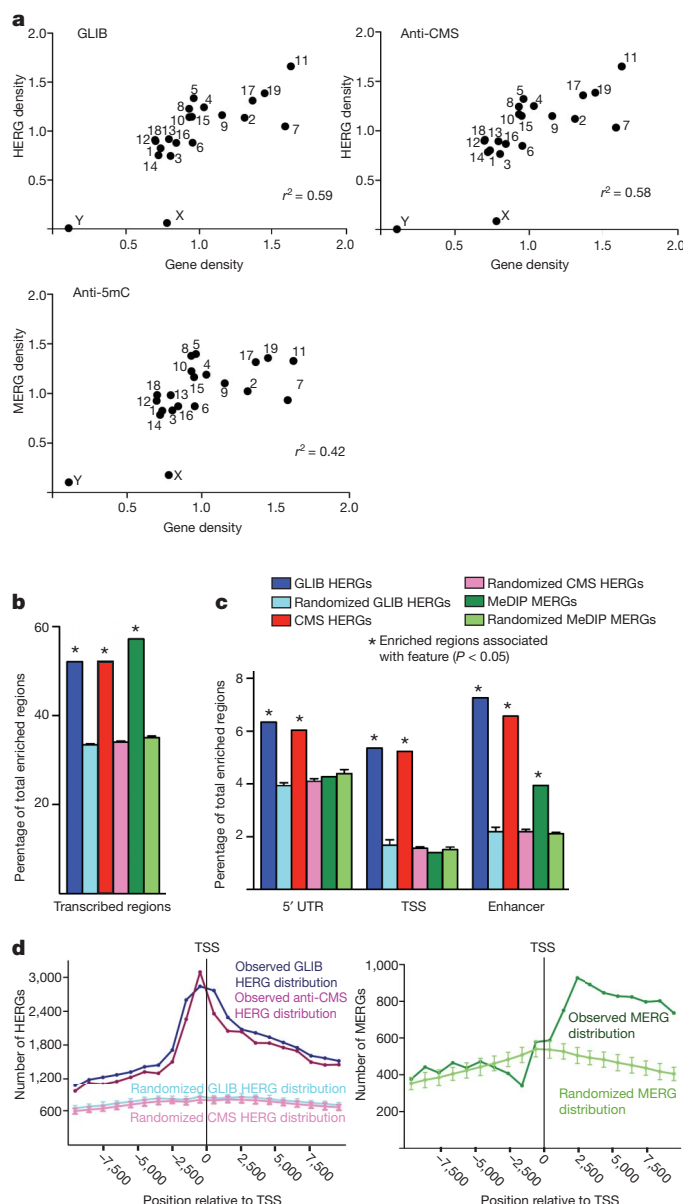


Figure 2 | Genomic distribution of 5hmC or 5mC enriched regions of the genome. **a**, Correlation of HERG or MERG density on each chromosome (y-axis) with gene density in the same chromosome (x-axis). Density is defined as frequency divided by chromosome length. **b, c**, Both HERGs and MERGs are enriched in transcribed regions (**b**), whereas HERGs are preferentially enriched at enhancers and the start sites of genes (**c**). The percentage of HERGs or MERGs mapping to the indicated genomic feature (darker bar) is compared with the percentage of randomly chosen sequences mapping to that feature (lighter bar). 5' UTR, 5' untranslated region. TSS, transcription start site (–800 bp to +200 bp relative to start of transcription). See Supplementary Methods for detailed definition of how HERGs or MERGs were classified as mapping to genomic features. **d**, Distribution of HERGs and MERGs relative to the TSS. The centre of each HERG was plotted relative to the nearest TSS in 1,000 bp increments from –10 kb to +10 kb surrounding the TSS.

HERGs and MERGs to the distribution of DNA fragments of equivalent length distributed randomly across the genome, both 5hmC and 5mC were enriched within transcribed regions, particularly exons, which are known to be sites of high CpG density¹⁷ as well as high DNA methylation¹⁸ (Fig. 2b and Supplementary Fig. 3c). However, only 5hmC was enriched at transcription start sites (TSSs) and within the 5' untranslated regions (UTRs) of genes (Fig. 2c). Moreover, 5hmC was relatively more enriched in enhancers (defined by H3K4me1 in the

absence of H3K4me3)¹⁹ than 5mC, strongly indicating a connection between 5hmC and regulatory elements (Fig. 2c). Plotting each HERG as a single point relative to the nearest TSS, we found that 5hmC is heavily enriched both 5' and 3' of the TSS, whereas 5mC is enriched primarily 3' of the TSS (Fig. 2d). These results show a unique distribution of 5hmC in regulatory elements of genes, one that is not explained simply by the distribution of 5mC, the substrate for TET enzymes.

The enrichment of 5hmC at the TSS suggested a role for 5hmC in transcriptional regulation. To evaluate this possibility, we used published data sets on gene expression^{20,21} and histone modification^{22,23} profiles in mouse ES cells to compare the sets of genes with 5hmC or 5mC at their start sites (Supplementary Tables 11–13) to the set of all genes in the genome. 5hmC is preferentially found at promoters with high or intermediate CpG content (Supplementary Fig. 4a), even though high CpG promoters are hypomethylated in ES cells^{16,18,24}. This distribution is consistent with the possibility that TET proteins are preferentially recruited to high CpG regions through their CpG-binding CXXC domains^{6,25}.

In ES cells, genes with 'bivalent' H3K27 and H3K4 trimethylation are transcriptionally inactive but poised for expression upon differentiation to embryoid bodies^{20,26,27}. We found that genes with 5hmC at their start sites were disproportionately likely to contain bivalent domains at their promoters; likewise, a majority (~60%) of genes reported to contain bivalent domains have 5hmC at their start sites (Fig. 3a). 5hmC was less likely to be found at genes with the activating 'H3K4me3 only' mark than is predicted by chance. Moreover, genes with 5hmC at their start sites showed lower expression in murine ES cells than other genes (Fig. 3b) and were more likely to be upregulated upon embryoid body differentiation (Fig. 3c). The correlation of 5hmC with bivalent domains holds even after adjusting for the known relation between promoter CpG content and bivalency²² (Supplementary Fig. 5). Although 5mC at the TSS also correlates with lower gene expression in murine ES cells (Supplementary Fig. 4b), 5mC is not enriched at the promoters of genes with bivalent domains²⁸ (Supplementary Fig. 4c), and genes with high levels of 5mC did not tend to be upregulated upon embryoid body differentiation (Supplementary Fig. 4d). Thus 5hmC is preferentially enriched at the promoters of genes with bivalent histone marks in ES cells, indicating that 5hmC may contribute functionally to the 'poised' but inactive state of these genes in ES cells.

Genes with 5hmC at their start sites are also disproportionately enriched in the set of genes whose promoters bind polycomb repressor complex (PRC) components, and in a majority of genes with the 'H3K27me3' only mark (Fig. 3a). There is a statistically significant correlation between genes that had 5hmC at the TSS and genes that were upregulated upon small interfering RNA-mediated Tet1 depletion⁸ (therefore, negatively regulated by Tet1) (Fig. 3d), indicating that 5hmC in the promoter region has a negative role in the transcription of some genes in ES cells. Unlike 5mC, however, 5hmC is not substantially enriched at sites of heterochromatic H3K9 or H4K20 trimethylation²² (data not shown).

Collectively, our results support a model in which 5hmC and 5mC have different roles in transcription. Like 5mC²⁸, 5hmC at promoters is predictive of lower levels of gene expression. However, 5hmC is uniquely associated with a 'poised' chromatin configuration and with genes that are upregulated upon differentiation, and may thus be involved in priming loci for rapid activation in response to appropriate signals. Activation of lineage-specific genetic loci upon differentiation could occur via a postulated 5mC 'demethylation' pathway (5mC to 5hmC to cytosine)¹ or through recruitment of transcriptional regulators that specifically recognize 5hmC and are activated in response to differentiation signals. The ability to profile 5hmC even at sparsely hydroxymethylated loci will allow a careful evaluation of these possibilities in differentiating cells.

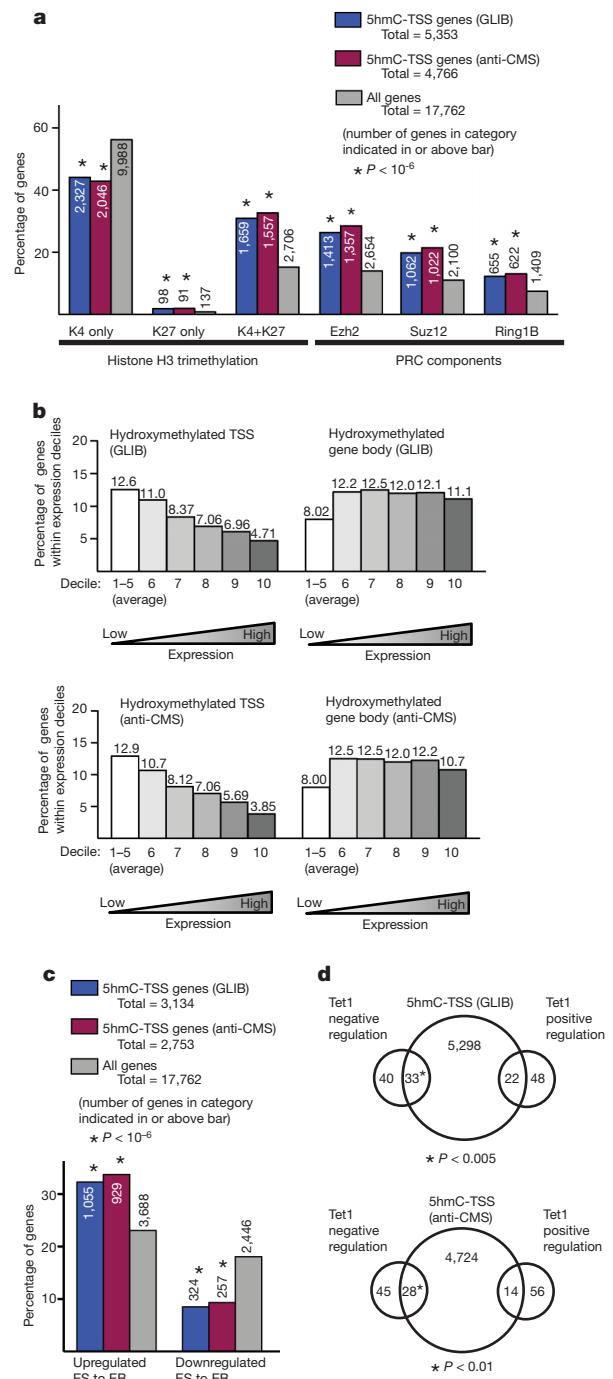


Figure 3 | Properties of HERGs at transcription start sites. **a**, The percentage of genes with 5hmC at the TSS (blue and red bars) reported to contain histone H3 trimethylation (left) or PRC components (right) at their promoters is compared to the fraction of all genes (grey bars) with these promoter marks²². Number of genes in each category is indicated. **b**, HERGs are enriched at the TSSs of genes with low expression in ES cells. All genes were ranked by level of expression in ES cells²¹ and sorted into deciles from lowest to highest. The per cent of genes within the decile category with 5-hmC enriched at the TSS (left) or within gene bodies (right) are shown for each methodology. The first five deciles, which are comprised of genes lacking statistically significant expression, are pooled and averaged in this analysis. **c**, HERGs are enriched at the TSS of genes upregulated upon differentiation to embryoid bodies (EB)²⁶. The percentage of genes with 5-hmC at their TSS (blue bars) that are substantially upregulated or downregulated upon differentiation to EB is compared with the percentage of total genes similarly regulated (grey bars). Number of genes in each category is indicated. **d**, Overlap between genes with 5hmC at the TSS and genes positively or negatively regulated by Tet1 (ref. 8).

METHODS SUMMARY

GLIB precipitation. V6.5 ES cells were lysed and proteins digested by treatment with Proteinase K at 55 °C. DNA was purified by phenol-chloroform extraction and then precipitated with ethanol. RNA was removed with RNase A (Qiagen). Samples were treated with 20 ng BGT per 1 µg DNA at 30 °C for 3 h (50 mM HEPES pH 8.0, 25 mM MgCl₂, 50 µM UDPG for 3 h at 30 °C), then oxidized with 23 mM sodium periodate 16 h at 22 °C in 0.1 M sodium phosphate pH 7.0. Periodate was quenched by the addition of 46 mM sodium sulphite at room temperature for 10 min, then exchanged into 1× PBS and incubated with 2 mM Aldehyde Reactive Probe (Invitrogen) for 1 h at 37 °C. DNA was sequenced with a HeliScope Single Molecule Sequencer. See Supplementary Methods for detailed protocol.

CMS precipitation. The generation of the anti-CMS antibody is described elsewhere⁵. DNA fragments were ligated with methylated adaptors and treated with sodium bisulphite (Qiagen). The DNA was then denatured for 10 min at 95 °C (0.4 M NaOH, 10 mM EDTA), neutralized by addition of cold 2 M ammonium acetate pH 7.0, incubated with anti-CMS antiserum in 1× immunoprecipitation buffer (10 mM sodium phosphate pH 7.0, 140 mM NaCl, 0.05% Triton X-100) for 2 h at 4 °C, and then precipitated with Protein G beads. Precipitated DNA was eluted with Proteinase K, purified by phenol-chloroform extraction, and amplified by 4–6 cycles PCR using Pfu TurboC_x hotstart DNA polymerase (Stratagene). DNA sequencing was carried out using Illumina/Solexa Genome Analyzer II and HiSeq sequencing systems.

Received 4 December 2010; accepted 11 April 2011.

Published online 8 May 2011.

- Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
- Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
- Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F. & Leonhardt, H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **38**, e181 (2010).
- Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
- Ko, M. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843 (2010).
- Iyer, L. M., Tahiliani, M., Rao, A. & Aravind, L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **8**, 1698–1710 (2009).
- Loenarz, C. & Schofield, C. J. Oxygenase catalyzed 5-methylcytosine hydroxylation. *Chem. Biol.* **16**, 580–583 (2009).
- Koh, K. P. *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
- Delhommeau, F. *et al.* Mutation in TET2 in myeloid cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
- Zhang, H., Li, X. J., Martin, D. B. & Aebersold, R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nature Biotechnol.* **21**, 660–666 (2003).
- Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotechnol.* **29**, 68–72 (2011).
- Hayatsu, H. & Shiragami, M. Reaction of bisulfite with the 5-hydroxymethyl group in pyrimidines and in phage DNAs. *Biochemistry* **18**, 632–637 (1979).
- Huang, Y. *et al.* The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE* **5**, e8888 (2010).
- Harris, T. D. *et al.* Single-molecule DNA sequencing of a viral genome. *Science* **320**, 106–109 (2008).
- Bowers, J. *et al.* Virtual terminator nucleotides for next-generation DNA sequencing. *Nature Methods* **6**, 593–595 (2009).
- Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
- Saxonov, S., Berg, P. & Brutlag, D. L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl Acad. Sci. USA* **103**, 1412–1417 (2006).
- Feng, S. *et al.* Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl Acad. Sci. USA* **107**, 8689–8694 (2010).
- Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.1016071107 (24 November 2010).
- Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
- Guttman, M. *et al.* *Ab initio* reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature Biotechnol.* **28**, 503–510 (2010).
- Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
- Ku, M. *et al.* Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* **4**, e1000242 (2008).
- Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
- Zhang, H. *et al.* TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res.* **20**, 1390–1393 (2010).
- Lee, T. I. *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301–313 (2006).
- Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
- Fouse, S. D. *et al.* Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2**, 160–169 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Ren for assistance in next generation sequencing using the Illumina platform. We thank M. Guttman for making his RNASeq data set available to us. W.A.P. is supported by a predoctoral graduate research fellowship from the National Science Foundation, and Y.H. by a postdoctoral fellowship from the Leukemia and Lymphoma Society. R.L. is supported by a California Institute for Regenerative Medicine Training Grant. This study was supported by the National Institute of Health grants RC1 DA028422, R01 AI44432 and 1 R01 HD065812-01A1 and a grant from the California Institute of Regenerative Medicine (to A.R.), a pilot grant from Harvard Catalyst, The Harvard Clinical and Translational Science Center (NIH Grant 1 UL1 RR 025758-02) and NIH K08 HL089150 (to S.A.), and a grant from the Mary. K. Chapman Foundation (to J.R.E.).

Author Contributions W.A.P., Y.B. and S.A. devised the GLIB method. W.A.P., S.A., H.R.H. and E.M.M. optimized the GLIB method. Y.H. generated the anti-CMS antiserum, and Y.H. and W.A.P. optimized the anti-CMS pull-down. W.A.P. and Y.H. grew ES cells. W.A.P. prepared GLIB samples for sequencing, Y.H. prepared CMS samples, H.R.H. performed MedIPs. Helicos sequencing and mapping was performed by P.K. and P.M.M., Illumina sequencing and mapping was performed by R.L. and J.R.E., and U.J.P. was responsible for bioinformatic analysis. M.K. performed the anti-5hmC dot blot. W.A.P. and M.T. performed anti-5hmC pull-downs. H.R.H. and S.M. performed and optimized *in vitro* tests of Tet substrate specificity. W.A.P., S.A. and A.R. wrote the manuscript. S.A. and A.R. coordinated research.

Author Information Data have been deposited at GEO under accession number GSE28682. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.A. (Suneet.Agarwal@childrens.harvard.edu) or A.R. (arao@idi.harvard.edu or arao@liai.org).

Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation

Gabriella Ficiz^{1*}, Miguel R. Branco^{1*}, Stefanie Seisenberger¹, Fátima Santos¹, Felix Krueger², Timothy A. Hore¹, C. Joana Marques^{1†}, Simon Andrews² & Wolf Reik^{1,3}

Methylation at the 5' position of cytosine in DNA has important roles in genome function and is dynamically reprogrammed during early embryonic and germ cell development¹. The mammalian genome also contains 5-hydroxymethylcytosine (5hmC), which seems to be generated by oxidation of 5-methylcytosine (5mC) by the TET family of enzymes that are highly expressed in embryonic stem (ES) cells^{2–4}. Here we use antibodies against 5hmC and 5mC together with high throughput sequencing to determine genome-wide patterns of methylation and hydroxymethylation in mouse wild-type and mutant ES cells and differentiating embryoid bodies. We find that 5hmC is mostly associated with euchromatin and that whereas 5mC is under-represented at gene promoters and CpG islands, 5hmC is enriched and is associated with increased transcriptional levels. Most, if not all, 5hmC in the genome depends on pre-existing 5mC and the balance between these two modifications is different between genomic regions. Knockdown of *Tet1* and *Tet2* causes downregulation of a group of genes that includes pluripotency-related genes (including *Esrrb*, *Prdm14*, *Dppa3*, *Klf2*, *Tcl1* and *Zfp42*) and a concomitant increase in methylation of their promoters, together with an increased propensity of ES cells for extraembryonic lineage differentiation. Declining levels of TETs during differentiation are associated with decreased hydroxymethylation levels at the promoters of ES cell-specific genes together with increased methylation and gene silencing. We propose that the balance between hydroxymethylation and methylation in the genome is inextricably linked with the balance between pluripotency and lineage commitment.

5hmC occurs in ES cells (in 5% of CpGs), Purkinje cells in the mouse brain, and in other adult mouse tissues^{2–5}. The TET1 and TET2 enzymes, which can oxidise 5mC thus generating 5hmC, are highly expressed in ES cells and regulate the expression of pluripotency-related genes together with the potential of ES cells to differentiate into the embryonic and extraembryonic lineages^{4,6}. The genomic distribution of 5hmC in the ES cell genome and during differentiation and its relation to the distribution to 5mC is unknown. Because bisulphite conversion and high throughput sequencing (BS-Seq) does not distinguish between 5mC and 5hmC⁷, we used specific antibodies (Fig. 1a and Supplementary Fig. 1) to determine the genomic distribution of both 5mC and 5hmC by MeDIP-Seq⁸ and hMeDIP-Seq (methylated DNA immunoprecipitation and hydroxymethylated DNA immunoprecipitation followed by high throughput sequencing, respectively) in two different ES cell lines (J1, E14), *Np95*^{−/−} ES cells (lacking maintenance methylation), *Tet1/2* knockdown ES cells, and embryoid bodies (EBs). We obtained 19–33 million paired-end reads for each sample; all samples were sequenced in two biological replicates which were found to be highly reproducible (Supplementary Table 1 and Supplementary Fig. 2). Note that (h)MeDIP-Seq profiles (as chromatin immunoprecipitation (ChIP)-Seq profiles) reveal only the relative distribution of the respective modification within a sample and therefore cannot be used to infer absolute quantitative differences between samples or antibodies.

By immunofluorescence we found strong nuclear staining for 5hmC in ES cells (and in other cell types) that broadly overlapped in euchromatic regions with staining for 5mC, whereas DAPI-dense heterochromatic regions are highly enriched for 5mC but not 5hmC (Fig. 1b and

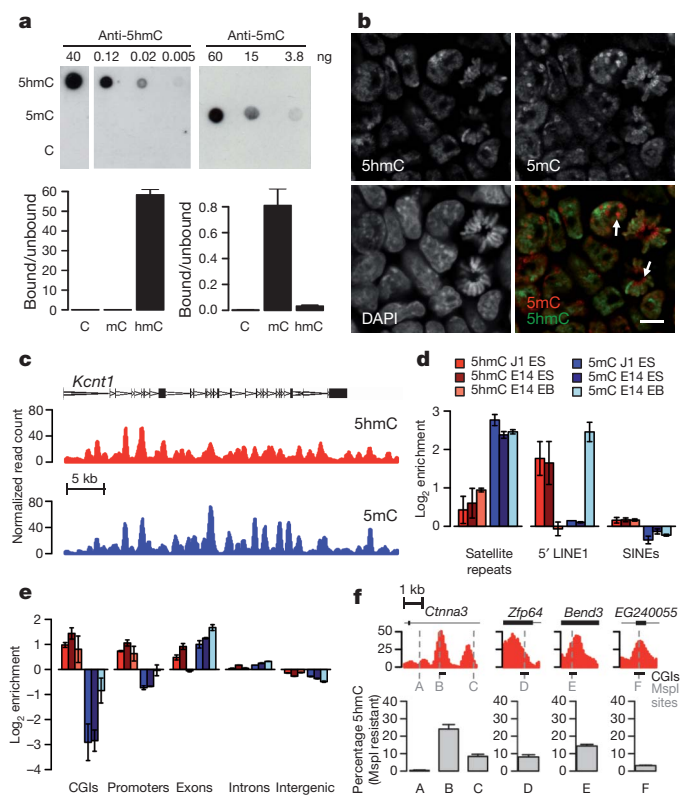


Figure 1 | Distribution of 5-hydroxymethylcytosine in the mouse genome.

a, The specificities of the antibodies used in this study were confirmed by dot blot and (h)MeDIP using PCR fragments containing 5hmC, 5mC or C.

b, Immunofluorescence co-staining of J1 ES cells with antibodies against 5hmC (green) and 5mC (red). Grey scale images of the two modifications are shown separately. Staining for 5mC is particularly strong in pericentromeric heterochromatin (arrows), contrary to 5hmC. Scale bar, 10 μm. **c**, Examples of hMeDIP-Seq and MeDIP-Seq profiles at a genomic region on Chr2 in J1 ES cells. **d**, Relative enrichment (log₂ bound/unbound) of 5hmC and 5mC in repetitive sequences in J1 and E14 ES cells and E14 EBs. **e**, Enrichment of 5hmC and 5mC in single-copy genomic features. Values in **d** and **e** represent means of two biological replicates with the ends of the error bars corresponding to the individual data points. **f**, Validation of the presence of 5hmC in CGIs using glucMS-qPCR (grey bars represent mean ± s.d.). Selected CGIs (black bars, upper panel) were tested for the presence of 5hmC at particular MspI sites (grey vertical line). Genomic coordinates of the left-most base pairs of each region: *Cttna3* (chr10, 63044495); *Zfp64* (chr2, 168750875); *Bend3* (chr10, 43230661); *EG240055* (also known as *Neur11b*: chr17, 26567975).

¹Laboratory of Developmental Genetics and Imprinting, The Babraham Institute, Cambridge CB22 3AT, UK. ²Bioinformatics Group, The Babraham Institute, Cambridge CB22 3AT, UK. ³Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK. †Present address: Genetics Department, Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal.

*These authors contributed equally to this work.

Supplementary Fig. 3). (h)MeDIP-Seq confirmed that 5hmC is widely distributed throughout non-repetitive regions (see example in Fig. 1c) and substantially overlaps with the distribution of 5mC, whereas satellite repeats (which are located in heterochromatin) are highly enriched for 5mC but substantially less for 5hmC (Fig. 1d and Supplementary Fig. 4). In single-copy regions the distribution of 5hmC in ES cells follows a broadly similar pattern to that of 5mC in intergenic regions, exons and introns, with a higher enrichment in exons over introns (Fig. 1e). Notably, whereas 5mC is relatively depleted from CpG islands (CGIs), gene promoters, the 5' ends of LINE1 elements (their promoters), CTCF and pluripotency transcription factor binding sites in accordance with previous work⁹, 5hmC is relatively enriched in all of these (Fig. 1d, e and Supplementary Figs 5 and 6). Furthermore, upon differentiation into EBs 5hmC enrichment decreases in these regions, concomitant with a gain of 5mC. Consistent with the distinct 5mC and 5hmC patterns at CGIs, whereas 5mC is depleted from high CpG density promoters (as described previously¹⁰), 5hmC remains enriched (Supplementary Fig. 7), indicating that the ratio of 5hmC to 5mC is higher here than in low CpG density promoters. To independently and quantitatively verify the presence of 5hmC in CGIs we carried out glucosylation of 5hmC in genomic DNA followed by MspI digestion (which does not digest glucosylated 5hmC) and quantitative PCR (qPCR) across MspI sites (glucMS-qPCR)¹¹. We found significant levels of 5hmC (3–24%) in selected CGIs (Fig. 1f and Supplementary Fig. 8). We also determined the corresponding 5mC levels (see Methods) and found them to be comparable to those of 5hmC at these CGIs, whereas elsewhere 5mC can be several fold higher than 5hmC (Supplementary Fig. 8, compare regions B and C). These measurements suggest that 5hmC is still derived from 5mC at CGIs, but that a high proportion of 5mC is converted to 5hmC in these regions. We also confirmed the presence of cytosine modifications in these regions by bisulphite conversion followed by Sequenom MassARRAY analysis (Supplementary Fig. 8).

We found by thin layer chromatography analysis that 5hmC was reduced in *Tet1/2* knockdown cells and *Np95*^{-/-} cells, and eliminated in *Dnmt1*^{-/-}/*Dnmt3a*^{-/-}/*Dnmt3b*^{-/-} triple knockout (TKO) ES cells (Fig. 2a). We confirmed this by glucMS-qPCR on selected regions and all were found to display lower levels of both modifications in *Np95*^{-/-} ES cells and only vestigial amounts in TKO cells (Fig. 2b). We find losses of 5hmC enrichment at exons, 5' regions of LINE1 elements and CTCF binding sites in *Np95*^{-/-} and *Tet1/2* knockdown ES cells (Fig. 2d and Supplementary Fig. 6). Enrichment of 5hmC is also reduced at promoters in *Np95*^{-/-} but not in *Tet1/2* knockdown cells. However, maintenance of relative enrichment levels at promoters upon *Tet1/2* knockdown means that its absolute 5hmC levels follow the observed genome-wide reduction in 5hmC (see Supplementary Fig. 9). Overall these results suggest that most 5hmC in mouse ES cells is dependent on pre-existing 5mC (although we cannot exclude that 5hmC may also be generated by an independent mechanism¹²), but that the kinetics of generating and maintaining 5mC and converting it to 5hmC are likely to be different for different genomic elements.

Our protocol for (h)MeDIP-Seq conserves information on strand-specificity of hydroxymethylation and methylation (Supplementary Fig. 10). Extensive occurrence of strand-biased regions was found in both 5mC and 5hmC methylomes, and these regions were enriched for CpH (where H is C, A, or T) dinucleotides (Fig. 3a), indicating that strand-specific (hydroxy)methylation occurs largely in non-CpG context. The strand specificity and sequence context of asymmetric (hydroxy)methylation were confirmed by analysis of a BS-Seq ES cell data set¹³ (Fig. 3c) and by bisulphite sequencing of selected asymmetric regions, showing that modification occurred predominantly in CpH context, where it was entirely strand-specific (Fig. 3d). Overall strand bias in 5hmC profiles was increased in *Np95*^{-/-} and *Tet1/2* knockdown cells (Fig. 3b), consistent with the dependence of 5hmC on pre-existing 5mC at CpGs and suggesting that TET1 and TET2 may have a preference for oxidizing 5mC in CpG context.

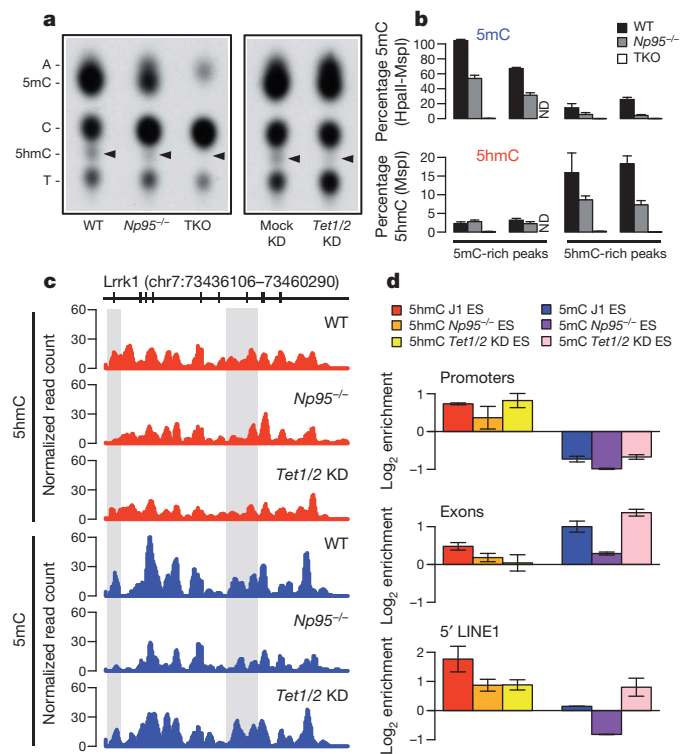


Figure 2 | Genetic relationship between methylation and hydroxymethylation. **a**, Thin layer chromatography separation of radioactively end-labelled bases from MspI-digested genomic DNA, showing reduced levels of 5hmC (arrowheads) in methylation- and TET-deficient ES cells. **b**, glucMS-qPCR validation of genomic regions specifically enriched for 5mC or 5hmC in wild-type (WT) J1, *Np95*^{-/-} and TKO ES cells (bars represent mean \pm s.d.). Genomic regions were selected on the basis of (h)MeDIP-Seq profiles of wild-type ES cells. **c**, Examples of (h)MeDIP-Seq profiles in wild type, *Np95*^{-/-} and *Tet1/2* KD ES cells. 5hmC profiles are relatively similar, whereas 5mC distribution is significantly altered in *Np95*^{-/-} cells, but less so in *Tet1/2* knockdown (*Tet1/2* KD) cells. Shaded areas highlight regions of altered 5hmC and/or 5mC enrichment. **d**, Relative enrichment at promoters, exons and 5' regions of LINE1 elements in J1, *Np95*^{-/-} and *Tet1/2* KD ES cells. *Np95* deficiency causes depletion of both 5hmC and 5mC in all three regions, whereas *Tet1/2* KD causes preferential reduction of 5hmC at exons and LINE1 promoters, which leads to increased 5mC enrichment in these regions. Values represent means of two biological replicates with the ends of the error bars corresponding to the individual data points.

We asked if there was a relationship between hydroxymethylation at gene promoters in ES cells and their transcription levels (Fig. 4a and Supplementary Fig. 11). We generated an ES cell transcriptome by RNA-Seq and classified promoters with respect to enrichment of 5mC and 5hmC. Whereas the presence of 5mC in the promoter region was associated with low levels of transcription as expected, 5hmC was associated with high levels of transcription. In fact, genes specifically enriched for 5hmC were more highly transcribed than those with neither of the modifications (Fig. 4a); this effect is also partially dependent on promoter CpG density (Supplementary Fig. 11). Promoters enriched for both 5hmC and 5mC were also associated with higher levels of transcription than promoters specifically enriched for 5mC, suggesting that presence of 5hmC partially overcomes the silencing effect of 5mC. Consistent with these observations, promoters that are high in 5hmC are enriched in the activating histone mark H3K4me3, whereas those enriched in 5mC are depleted of H3K4me3 (Fig. 4b; data from ref. 14). 5hmC in exons was also found associated with increased levels of transcription (Supplementary Fig. 11), consistent with what has been found in mouse cerebellum¹⁵.

RNA-Seq of *Tet1/2* knockdown ES cells identified 107 genes that were downregulated in knockdown cells (18 out of 22 validated by

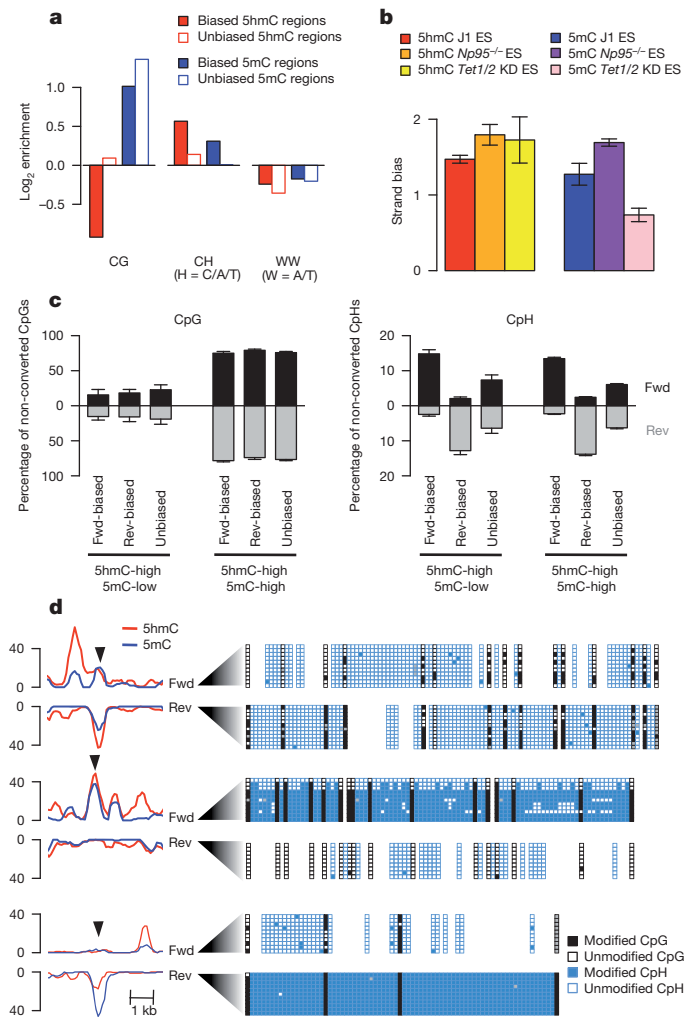


Figure 3 | Strand specificity and sequence context of methylation and hydroxymethylation. **a**, Enrichment of dinucleotide sequences present in the central 200 bp of 5hmC and 5mC regions separated into biased and unbiased fragments. CG dinucleotides are enriched in unbiased regions, as expected from its symmetric nature. Biased regions are enriched for CH dinucleotides, indicating extensive non-CpG (hydroxy)methylation. **b**, Strand bias measurements, which represent the overall level of asymmetric methylation in the genome. Depletion of NP95 increases strand bias in the 5hmC and 5mC profiles due to reduced CpG methylation. Knockdown of *Tet1/2* decreases strand bias in the 5mC profile, as expected if reduction of 5hmC at CpGs leads to an accumulation of 5mC at the same sites. Values represent means of two biological replicates with the ends of the error bars corresponding to the individual data points. **c**, BS-Seq¹³ validation of the (h)MeDIP data. Percentages of methylated CpGs present in 5hmC-enriched peaks containing low or high 5mC levels are plotted (left) showing the symmetrical nature of CpG methylation (in both biased and unbiased peaks). Conversely, CpH methylation in biased peaks is asymmetric in nature (right). Error bars represent 95% confidence intervals. **d**, Validation of asymmetric methylation by bisulphite sequencing of biased and unbiased (h)MeDIP-Seq peaks. Filled squares represent methylated/hydroxymethylated cytosines and empty squares represent unmethylated cytosines. Bisulphite sequencing confirms the asymmetric nature of methylation in biased regions (middle and bottom) and reveals extensive non-CpG methylation, whereas the unbiased region (top) contains mostly CpG methylation.

quantitative PCR with reverse transcription (qRT-PCR); Fig. 4c and Supplementary Fig. 12). We also carried out *Tet1* knockdown on its own and found all 18 validated genes consistently downregulated by qRT-PCR (Supplementary Fig. 12), indicating that TET1 has the major role in the observed expression changes. This was confirmed in a stable ES cell line containing a doxycycline-inducible shRNA

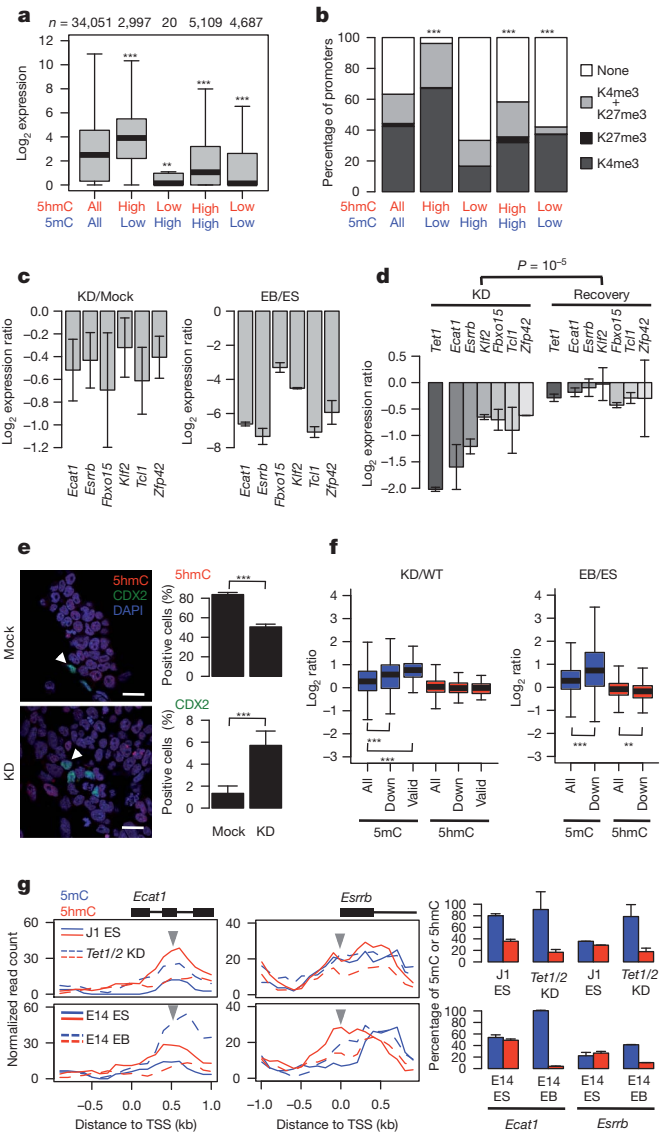


Figure 4 | Gene expression and promoter methylation in ES cells and during differentiation. **a**, Relationship between 5hmC and 5mC levels at gene promoters and expression of downstream genes measured by RNA-Seq in J1 ES cells. Significance levels are relative to all promoters (**P < 0.001, ***P < 0.0001 throughout the figure). **b**, Relationship between 5hmC, 5mC and presence/absence of H3K4me3 and H3K27me3 at gene promoters (data from ref. 14). **c**, qRT-PCR validation of genes downregulated upon *Tet1/2* KD across three biological replicates and expression level changes in the same genes from ES to EB differentiation (values are mean ± s.d.). **d**, Induction of a shRNA targeting *Tet1* in a stable ES cell line also results in the downregulation of the genes in **c**. Restoring TET1 expression leads to recovery in expression of these genes (values are mean ± s.d.). **e**, Co-immunostaining of mock and *Tet1/2* KD ES cells for 5hmC (red) and CDX2 (green). Scale bar, 20 μm. Cells were scored for presence of 5hmC and CDX2 expression (n = 1,120 and 1,209 for mock and *Tet1/2* KD cells, respectively; values are percentage of cells ± 95% confidence interval). **f**, KD/WT ratios for promoter 5mC and 5hmC shows that downregulated genes, and in particular qRT-PCR-validated ones, suffer methylation changes different from the pool of all genes. Genes downregulated upon ES to EB differentiation have increased 5mC enrichment levels and decreased 5hmC. **g**, Examples of 5hmC and 5mC profiles of two genes in J1 and E14 ES cells, *Tet1/2* KD and E14 EBs, and corresponding quantification of 5mC and 5hmC levels by glucMS-qPCR (the *MspI* site used is indicated by the grey arrowhead). A considerable reduction in 5hmC levels is detected upon both *Tet1/2* KD and differentiation into EBs, with a concomitant increase in 5mC.

against *Tet1* in which we found again downregulation of the same set of genes (Fig. 4d). Removal of doxycycline led to recovery of TET1 and 5hmC levels back to normal and we found that gene expression changes were restored to wild-type levels (Fig. 4d and Supplementary Fig. 13).

The genes that were downregulated in response to *Tet1/2* knockdown included pluripotency-related genes such as *Esrrb*, *Klf2*, *Tcl1*, *Zfp42*, *Dppa3*, *Ecat1* (also known as *2410004A20Rik*) and *Prdm14* (by contrast *Nanog*, *Oct4* (also known as *Pou5f1*) and *Sox2* were not downregulated). These are amongst the earliest genes to be downregulated upon ES cell differentiation¹⁶, and include genes that undergo epigenetically regulated transcriptional fluctuations in ES cells, such as *Dppa3* and *Zfp42* (refs 17, 18). It remains to be seen whether such fluctuations are associated with stochastic patterning of 5hmC at these promoters across a cell population. Consistent with a role in regulating transcription of pluripotency-associated genes, the upstream region of the *Tet1* gene contains a large cluster of binding sites for core pluripotency transcription factors (Supplementary Fig. 14), and both TET1 and TET2 are rapidly downregulated upon differentiation to EBs (Supplementary Fig. 15). Indeed, we found a 23-fold enrichment ($P < 2.2 \times 10^{-16}$) for transcripts downregulated upon *Tet1/2* knockdown to also be substantially downregulated during ES cell differentiation¹⁹, 16 of which we validated by qRT-PCR (Fig. 4c and Supplementary Fig. 12).

Although ES cells subjected to *Tet1/2* knockdown did not appear to differentiate spontaneously, we found that markers of extraembryonic endoderm differentiation (*Gata6* and *Sox17*) were precociously expressed when *Tet1/2* knockdown ES cells were differentiated with retinoic acid (Supplementary Fig. 16). This is consistent with downregulation of *Esrrb* and *Prdm14*, which safeguard ES cells from commitment to endoderm cell fate^{20,21}. We also found a significant increase of CDX2-positive cells concomitant with the decrease in the number of 5hmC-positive cells upon *Tet1/2* knockdown (all CDX2-positive cells were particularly low in 5hmC, Fig. 4e). These results are in agreement with recent studies using *Tet1* knockdown ES cells^{4,6}.

Promoters of genes that were downregulated during ES cell differentiation had a marked decrease of 5hmC enrichment levels, which was accompanied by a significant increase in 5mC levels (Fig. 4f, g). Notably, genes that were downregulated in *Tet1/2* knockdown cells had robustly increased levels of 5mC in their promoters (Fig. 4f, g; see also Supplementary Fig. 17), suggesting that decline of TET levels during differentiation leads, at least in part, to gene silencing through methylation of promoters. Enrichment of 5hmC in these promoters was unchanged, that is, its absolute levels accompanied the genome-wide loss of 5hmC. Importantly, we confirmed the reduction of 5hmC levels by glucMS-qPCR of selected promoters (Fig. 4g).

Our study shows that 5hmC is relatively enriched in euchromatic parts of the genome, including in CGIs and promoters, and its presence in promoters and exons is associated with increased levels of transcription. This may in part be explained by the removal of the repressive effects of 5mC, but it is also possible that 5hmC itself has a positive effect on transcription. Furthermore, the TET proteins may have roles in transcriptional regulation in addition to their ability to convert 5mC to 5hmC. The core pluripotency network is connected with TET1/2 which regulate genes with established roles in pluripotency and epigenetic reprogramming such as *Esrrb*, *Dppa3*, *Klf2*, *Zfp42* and *Prdm14*, thus safeguarding ES cells against commitment to extraembryonic cell fate (Supplementary Fig. 19). On the basis of these results, we suggest that hydroxymethylation and the TET proteins could also have a role in erasing methylation marks from promoters of pluripotency-related genes during fusion of ES cells with somatic cells^{22,23}, and during the generation of induced pluripotent stem cells for which erasure of DNA methylation seems critical²⁴. Hydroxymethylation may also have a role in the large-scale erasure of methylation in primordial germ cells and early embryos^{25,26}.

METHODS SUMMARY

All immunofluorescence and (h)MeDIP-Seq data were produced using a rabbit anti-5hmC polyclonal antibody (Active Motif, catalogue no. 39769) and a mouse

anti-5mC monoclonal antibody (Eurogentec, MMS-900P-B). J1, E14, *Np95*^{-/-} and TKO ES cells were grown under standard conditions in the presence of serum and leukaemia inhibitory factor (LIF). Differentiation of E14 ES cells into embryoid bodies was done by removal of LIF and suspension culture for 13 days. RNA interference (RNAi) experiments were performed as described² with modifications. A2lox.cre ES cells²⁷ were targeted with a short hairpin RNA-micro RNA (shRNA-mir) sequence against *Tet1* cloned into p2Lox; induction of *Tet1* knockdown and recovery were achieved by the addition and later removal of doxycycline. Cells were either fixed for immunofluorescence or collected for DNA and/or RNA extraction. MeDIP-Seq and hMeDIP-Seq were based on the MeDIP method²⁸ using either the anti-5mC or anti-5hmC antibodies, respectively, but incorporating the ligation of Illumina adaptors for paired-end sequencing, which was performed on a Illumina Genome Analyzer GAII-X. RNA-Seq was performed as described previously²⁹ with modifications. Bioinformatic analyses were performed using SeqMonk (<http://www.bioinformatics.bbsrc.ac.uk/projects/seq-monk/>) and custom Perl or R scripts.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 June 2010; accepted 18 March 2011.

Published online 3 April 2011.

- Feng, S., Jacobsen, S. E. & Reik, W. Epigenetic reprogramming in plant and animal development. *Science* **330**, 622–627 (2010).
- Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
- Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
- Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
- Szwagierczak, A. *et al.* Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **38**, e181 (2010).
- Koh, K. P. *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
- Huang, Y. *et al.* The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE* **5**, e8888 (2010).
- Down, T. *et al.* A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nature Biotechnol.* **26**, 779–785 (2008).
- Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
- Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genet.* **39**, 457–466 (2007).
- Davis, T. & Vaisvila, R. High sensitivity 5-hydroxymethylcytosine detection in Balb/C brain tissue. *J. Vis. Exp.* **48**, doi:10.3791/2661 (2011).
- Liutkeviciute, Z. *et al.* Cytosine-5-methyltransferases add aldehydes to DNA. *Nature Chem. Biol.* **5**, 400–402 (2009).
- Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105 (2010).
- Ku, M. *et al.* Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* **4**, e1000242 (2008).
- Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotechnol.* **29**, 68–72 (2011).
- Glover, C. H. *et al.* Meta-analysis of differentiating mouse embryonic stem cell gene expression kinetics reveals early change of a small gene set. *PLoS Comput. Biol.* **2**, e158 (2006).
- Hayashi, K., Lopes, S. M., Tang, F. & Surani, M. A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401 (2008).
- Toyooka, Y. *et al.* Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135**, 909–918 (2008).
- Cloonan, N. *et al.* Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods* **5**, 613–619 (2008).
- Ivanova, N. *et al.* Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533–538 (2006).
- Ma, Z. *et al.* Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. *Nature Struct. Mol. Biol.* **18**, 120–127 (2011).
- Tada, M. *et al.* Nuclear reprogramming of somatic cells by *in vitro* hybridization with ES cells. *Curr. Biol.* **11**, 1553–1558 (2001).
- Bhutani, N. *et al.* Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042–1047 (2010).
- Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
- Iqbal, K., Jin, S. G., Pfeifer, G. P. & Szabo, P. E. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl Acad. Sci. USA* **108**, 3642–3647 (2011).
- Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Commun* **2**, 241 (2011).
- Mallanna, S. K. *et al.* Proteomic analysis of Sox2-associated proteins during early stages of mouse embryonic stem cell differentiation identifies Sox21 as a novel regulator of stem cell fate. *Stem Cells* **28**, 1715–1727 (2010).

28. Weber, M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genet.* **37**, 853–862 (2005).
29. Marioni, J. C. *et al.* RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **18**, 1509–1517 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank A. Bird for discussions and input, K. Tabbada for help with epigenomics and transcriptomics sequencing, and M. Okano for providing the TKO ES cells. We would like to thank H. Stuart for her dedicated contribution to gene expression analyses, and all Reik lab members for advice and discussion. This work was funded by the BBSRC, MRC and the EU NoE The Epigenome.

Author Contributions G.F. and M.R.B. designed and performed experiments and analysed data. S.S. established the (h)MeDIP-Seq protocol and performed bisulphite sequencing. F.S. performed immunostainings. T.A.H. carried out qRT-PCR and glucMS-qPCR analyses. C.J.M. established the inducible *Tet1* shRNA ES cell line. F.K. and S.A. performed bioinformatic analyses. W.R. designed and directed the study. G.F., M.R.B. and W.R. wrote the manuscript.

Author Information All sequencing files have been deposited at the EBI Sequence Read Archive under the accession number ERP000570 (<http://www.ebi.ac.uk/ena/data/view/ERP000570>). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to W.R. (wolf.reik@bbsrc.ac.uk).

METHODS

Antibody validation by dot blot and methylated/hydroxymethylated DNA immunoprecipitation assay (MeDIP/hMeDIP). All immunofluorescence and (h)MeDIP-Seq data were produced using a rabbit anti-5hmC polyclonal antibody (Active Motif, catalogue no. 39769) and a mouse anti-5mC monoclonal antibody (Eurogentec, MMS-900P-B). To generate control templates, PCR fragments were amplified from M13mp18 DNA or custom oligonucleotides using either dCTP, d5mCTP or d5hmCTP. For dot blot, analysis DNA samples were denatured at 99 °C for 5 min and spotted onto Hybond-N+ nitrocellulose membranes (GE Healthcare). After ultraviolet cross-linking membranes were blocked overnight with 10% non-fat milk and 1% BSA in PBT (PBS + 0.1% Tween20) at 4 °C followed by >1 h incubation with either the 5mC or 5hmC antibodies (1:500 in blocking solution) at room temperature. Membranes were washed four times with PBT, incubated for 30 min with horseradish peroxidase (HRP)-conjugated goat anti-mouse or anti-rat antibodies (GE Healthcare; 1:10,000 in blocking solution), washed with PBT, and developed using the ECL+ detection system (GE Healthcare). For (h)MeDIP, three control templates with different sequences (~200-bp products containing C, 5mC or 5hmC; 15 pg each) were mixed with sonicated genomic DNA (1.5 µg) followed by denaturation (10 min at 95 °C) and immunoprecipitation as described previously for MeDIP²⁸ using 2 µg of anti-5mC or anti-5hmC antibody and 8 µl Dynabeads (coupled with M-280 sheep anti-mouse IgG for the 5mC antibody or with Protein G for the 5hmC antibody, Invitrogen). Pulled-down products were detected by qPCR and normalized to the unbound fraction.

Cell lines and other biological samples. J1 ES cell line (129S4/SvJae) was purchased from ATCC (catalogue no. SCRC-1010) and grown on a γ -irradiated pMEF feeder layer at 37 °C and 5% CO₂ in complete ES medium (DMEM 4,500 mg l⁻¹ glucose, 4 mM L-glutamine and 110 mg l⁻¹ sodium pyruvate, 15% fetal bovine serum, 100 U of penicillin/100 µg of streptomycin in 100 ml medium, 0.1 mM non-essential amino acids, 50 µM β -mercaptoethanol, 10³ U LIF ESGRO). E14 ES cells were grown either in complete ES medium or differentiated for 13 days into embryoid bodies via LIF removal and suspension culture. Retinoic acid (RA) differentiation of the mock siRNA treated and *Tet1/2* knockdown ES cells was done with 2 µM end concentration of RA in complete ES medium without LIF for 24 h. *Np95*^{-/-} ES cells (129/Ola derived) and TKO ES cells (*Dnmt3a*^{-/-}, *Dnmt3b*^{-/-} and *Dnmt1*^{-/-}, J1-derived, gift from M. Okano³⁰) were grown in complete ES medium. Primary mouse embryonic fibroblasts were derived from E11.5 embryos (B6CBAF1 \times B6) and grown for three passages in DMEM 4,500 mg l⁻¹ glucose, 4 mM L-glutamine and 110 mg l⁻¹ sodium pyruvate, 10% fetal bovine serum, 100 U of penicillin/100 µg of streptomycin in 100 ml medium, 50 µM β -mercaptoethanol. Cerebellum sections were a gift from E. Ivanova and G. Kelsey.

RNAi knockdown of *Tet1* and *Tet2* in ES cells. RNA interference experiments were performed as described² with modifications. Transfections of Dharmacon siGENOME siRNA duplexes (Thermo Fisher Scientific) against mouse *Tet1* (catalogue no. D-062861-01; caacuugcauccagauua), siGENOME SMARTpool against *Tet2* (catalogue no. M-058965-01; gaaagcagcucgaaagcg, ccucagauuuuaggaga, acuaacacuccaccuua, uagcaacguuuucuccuua) and siGENOME non-targeting siRNA#2 (catalogue no. D-001210-02; sequence not available) were done with Lipofectamine 2000 according to the manufacturer's instructions. Cells were harvested after three rounds of transfection for DNA/RNA isolation.

Stable shRNA *Tet1* knockdown and recovery. The pSM2 retroviral vector containing the shRNA-mir sequence targeting the *Tet1* messenger RNA (tgctgttgac agtgagcgcgtagctatagatagatagatagtggaagccacgatgtattactactctatagctagctctactgc ctgga) was purchased from Open Biosystems. shRNA-mir sequences were amplified by PCR using primers that created restriction sites for HindIII and NotI (atacaagcttcagggaattgttggaagggc and agcgccgcgctcttcaattgaaaaaagtga), cloned into p2Lox³¹ and subsequently transfected into A2lox.cre ES cells²⁷ (derived from the E14 cell line strain 129P2/OlaHsd, provided by M. Kyba). One day before transfection, Cre expression was induced by adding doxycycline (0.5 µg ml⁻¹) to the complete ES cell medium to promote the stable site-specific integration of the shRNA-mir sequence, following inducible cassette exchange recombination, which renders the cells resistant to neomycin.

ES cells were transfected using Lipofectamine 2000 (Invitrogen) at a concentration of 5×10^5 cells ml⁻¹. One day after transfection, selection medium containing geneticin (G418, Melford; 300 µg ml⁻¹ active concentration) was added to the cells and selection was maintained for 10 days. Resistant colonies were then individually picked into 96-well plates and expanded for freezing of stable cell lines. Integration was confirmed by PCR using Loxin primers³². shRNA-mir expression was induced by adding 2 µg ml⁻¹ of doxycycline (Sigma) to the culture medium during 5 days followed by removal of the doxycycline from the ES medium for an additional 7 days. RNA and DNA were isolated before removing the doxycycline to evaluate

knockdown effects and after 7 days of recovery, using a Qiagen AllPrep DNA/RNA isolation kit.

Immunofluorescence, microscopy and image analysis. Antibody staining of DNA methylation and hydroxymethylation was performed as previously described³³ with modifications. Briefly, cells were fixed with 4% PFA for 15 min and, after permeabilization with 0.5% Triton X-100, the samples were treated with 4 N HCl for 10 min at room temperature, washed in PBS Tween and blocked overnight; simultaneous incubation with both primary antibodies followed by simultaneous secondary detection was used. Antibody staining against CDX2 was performed as previously described³⁴. Mouse cerebellum cryosections (30 µm) were fixed with methanol (20 min at -20 °C) before staining with anti-calbindin D28K (CBP; gift from P. Emson) and post-fixing with 2% PFA after which same protocol as above was used for 5hmC staining. Single optical sections were captured with a Zeiss LSM510 Meta microscope ($\times 63$ oil-immersion objective) and the images pseudo-coloured using Adobe Photoshop. Semi-quantification of signals was performed on single optical sections using Velocity5.2 (Improvision).

DNA/RNA extraction. Genomic DNA was prepared using the Qiagen AllPrep DNA/RNA mini kit. RNA was extracted using either the Qiagen AllPrep DNA/RNA mini kit or RNeasy mini kit and subjected to DNase treatment using the Ambion DNA-free kit according to the manufacturers' instructions.

(h)MeDIP and next generation sequencing. The (h)MeDIP-Seq protocol was performed as described above with the following modifications: after sonication of gDNA the ends of the DNA fragments were repaired and paired-end sequencing specific adaptors (Illumina) were ligated using either a Paired-End DNA Sample Preparation Kit (Illumina) or NEBNext DNA Sample Prep Reagent Set 1 (NEB). Following adaptor ligation, DNA was immunoprecipitated and purified. Fragments were amplified with 12–18 cycles using adaptor specific primers (Illumina); fragments ranging between 300 and 500 bp in size were gel-purified before cluster generation and sequencing. Sequencing was done on an Illumina Genome Analyzer GAIIX using Cluster Generation v2 and 4 chemistries as well as Sequencing by Synthesis Kits v3 and v4. Data collection was performed using Sequencing Control Software v2.5 and 2.6. Real-time Analysis (RTA) 1.5–1.8 were used for base calling. Genomic mapping of short reads was performed using the sequence_pair mode of ELAND in the Illumina CASSAVA pipeline v1.5–1.8. Details on the number of sequencing reads obtained for each run are shown in Supplementary Table 1.

Bisulphite sequencing. J1 ES cell genomic DNA was bisulphite-treated using the Qiagen Epitect Kit and amplified using either Qiagen Hotstar Taq DNA Polymerase or Roche High Fidelity DNA Polymerase (primer sequences in Supplementary Table 2). A single amplification band was excised from the agarose gel and cloned into pGEM-T Easy vector (Promega) for sequencing.

Sequenom MassARRAY. Genomic DNA was bisulphite treated using the Qiagen Epitect Kit. PCR amplification of target regions (primer sequences in Supplementary Table 2), *in vitro* transcription and cleavage of the products for MassARRAY analysis were performed according to the manufacturer's instructions.

Glucosylation of genomic 5hmC followed by methylation sensitive qPCR (glucMS-qPCR). Genomic DNA (1 µg) was treated with T4 Phage β -glucosyl-transferase (T4-BGT, NEB M0357S) according to the manufacturer's instructions. Glucosylated genomic DNA (100 ng) was digested with 10 U of either HpaII, MspI or no enzyme (mock digestion) at 37 °C overnight, followed by inactivation for 20 min at 80 °C. The HpaII- and MspI-resistant fraction was quantified by qPCR using primers designed around at least one HpaII/MspI site, and normalizing to the mock digestion control and two regions lacking HpaII/MspI sites (Supplementary Table 2). Resistance to MspI directly translates into percentage of 5hmC, whereas 5mC levels were obtained by subtracting the 5hmC contribution from the total HpaII resistance¹¹.

5hmC detection using thin layer chromatography (TLC). Detection of 5hmC within MspI sites was done as described previously². Briefly, 1 µg of genomic DNA was restriction-enzyme-digested with 20 U MspI and 10 µg RNase A overnight at 37 °C, followed by inactivation of the enzyme at 65 °C for 20 min. DNA fragments were dephosphorylated with shrimp alkaline phosphatase and purified using QIAquick PCR purification kit followed by radioactive end labelling with ³²P-ATP (10 µCi, 3.3 pmol) using T4 Polynucleotide kinase for 1 h at 37 °C. Radioactively labelled DNA was precipitated, resuspended in 18 µl DNase I buffer and fragmented to single nucleotides with 1 µl DNase I (10 U µl⁻¹; Roche) and 1 µl SVDP (10 µg µl⁻¹; Worthington) for 3 h at 37 °C. Samples of 1–5 µl were spotted onto PEI cellulose F TLC plates and developed with isobutyric acid:H₂O:NH₃ (66:20:1 v/v/v) overnight followed by drying of the plate and exposure of radioactivity on an imaging film.

mRNA library preparation for next generation sequencing (RNA-Seq). mRNA was isolated from 3 µg total RNA using Dynabeads mRNA DIRECT (Invitrogen) and fragmented with RNA fragmentation reagent (Ambion). First strand cDNA synthesis was done with SuperScript III First-Strand Synthesis System and 3 µg µl⁻¹

random hexamers (Invitrogen) followed by second strand synthesis with DNA Polymerase I and RNase H. After purification, a sequencing library was generated from the double stranded cDNA using paired-end adaptors (Illumina) and NEBNext DNA Sample Prep Reagent Set 1 (NEB), and sequenced following a single-end sequencing protocol²⁹. Sequencing was done on an Illumina Genome Analyzer GAII-X using Cluster Generation v4 chemistry and Sequencing by Synthesis Kit v4. Data collection was performed using Sequencing Control Software v2.6. Real-time Analysis (RTA) 1.6 was used for data monitoring. Spliced mapping of RNA-Seq data was performed with TopHat v1.0.14 using default parameters³⁵. Details on the number of sequencing reads obtained for each run are shown in Supplementary Table 1.

Quantitative reverse transcription PCR. RNA was extracted using the Qiagen RNeasy Mini kit and subjected to DNase treatment using the Ambion DNA-free kit according to the manufacturers' instructions. cDNA was constructed from 2 µg of this RNA using the SuperScript III First-Strand Synthesis System for RT-PCR using random hexamers to prime the reaction. This cDNA was diluted 1:50 and used as template for quantitative real-time PCR reactions in combination with Brilliant II SYBR Green QPCR Master Mix and primers designed to specifically amplify a small product (intron-spanning where possible) for each gene of interest (Supplementary Table 2). Cycling reactions were performed in duplicate and cycle threshold (C_t) fluorescence data recorded on a Stratagene Mx3000P thermal cycler and Bio-Rad C1000 Thermal Cycler. The relative abundance of each gene of interest was calculated on the basis of the $\Delta\Delta C_t$ method³⁶, where results were normalized to the average C_t of two housekeeping genes with consistent C_t values over all samples (*Atp5b* and *Hsp90ab1*).

Bioinformatics. Analysis of gene region enrichment. An initial identification of enriched clusters was performed on a data group, which combined the reads from all of the individual data sets. Clusters were identified where the density of reads represented an enrichment of >1 fold (equating to 19 reads) over at least 50 bp. Where adjacent clusters were within 20 bp of each other they were combined. All clusters were then quantified with \log_2 -transformed count of the number of overlapping reads. The read counts were normalized both to the total count of the largest data set by applying a linear scaling factor, and by the length of the cluster. Clusters with a quantified value of more than 12 in any sample were removed because these represented an unrealistically high level of enrichment which was most likely due to mismapping of data. Each cluster was then called as present or absent in each sample. A cutoff of 5 (approximately equivalent to the median in most samples) in the corrected counts was taken as the point above which a cluster would be said to be present. All clusters were divided into groups based on the position of the centre of the cluster falling into a promoter, gene, exon, intron or intergenic region. Finally counts were made for the number of present clusters per kilobase of sequence in each of the different genomic region classes. Final enrichment values were calculated as the \log_2 ratio of the clusters per kb in the selected genomic region compared to the whole genome cluster density. Similar results were obtained when normalizing the data to an unbound fraction of a MeDIP experiment subjected to sequencing, ruling out any potential mapping effects.

Strand bias analysis. Clusters were generated as for the region enrichment analysis. For each cluster a count was made for the number of forward and reverse reads overlapping with that cluster in every sample. A cluster was discarded in a sample if the total count of forward plus reverse was less than 20 or greater than 200. For each valid cluster a bias value was calculated as $\text{abs}(\log_2((f+1)/(r+1))) - c(f+r)$ where $c(f+r)$ represents an averaged value from a simulation of observed strand biases from clusters with different numbers of reads, but where the probability of reads being either forward or reverse was exactly 0.5. The valid bias values were then attributed to whichever genomic regions they fell into and an average value for each genomic region for each sample was calculated. High values would indicate a higher than expected level of asymmetry in individual clusters and a zero value would

indicate a completely symmetrical sample. A control sample of sonicated input DNA showed a mean bias of +0.021 (data not shown).

Repeat analysis. For non-directional repeat analysis all currently known instances of repeats were retrieved from the Ensembl database (totalling nearly 9.5 million), and the sequence information for the major types of repeats were concatenated to form individual repeat genomes. The repeat content of different samples was determined using in-house developed software that aligns all sequences of Illumina sequence files to the entire repeatome employing multiple instances of Bowtie³⁷. The number of aligning sequences was counted for each repeat type individually.

RNA-Seq analysis. Initial RNA-Seq quantification was performed over each exon and was expressed as the number of overlapping bases of sequence per base of exon. The values were normalized to total read count between samples and an overall expression level was calculated for each transcript by normalizing the total expression value of the exons in that transcript to its total exon length. On the *Tet1/2* knockdown experiment we found that effects in expression were proportional to the amount of measured knockdown, and the first of three biological replicates had considerable lower amounts of *Tet1* and *Tet2* transcripts when compared with the remaining. We therefore classified a transcript as differentially expressed if: (1) its expression level was above 2.5; (2) it was up- or downregulated in the first replicate by >1.5-fold relative to both untransfected and mock KD controls; (3) it was up- or downregulated by >1.15-fold in the other two replicates. Despite these low thresholds 18 out of 22 downregulated genes were validated by qRT-PCR in five biological replicates. For the comparison of the published ES and EB data sets¹⁹, we classified transcripts as differentially expressed if their expression ratio was larger than fourfold. For correlation of expression with 5mC and 5hmC levels, promoters were classified as 5mC- or 5hmC-high if their normalized read count (\log_2) was above 5.5 and 5mC- or 5hmC-low if it was below 4.5 (medians were between 4.2 and 5.5). Similar results were obtained with a large range of threshold values or by using quantile-based thresholds.

BS-Seq analysis. A mouse ES cell shotgun BS-Seq data set¹³ was obtained from GEO (accession number GSE19960) and remapped to the mouse genome (build NCBI37) using Bismark (<http://www.bioinformatics.bbsrc.ac.uk/projects/bismark>). Prior to performing alignments the first 5 bp of all reads in the mouse ES cell shotgun data were clipped off to remove adaptor sequence. To validate our results with the above mentioned publically available BS-Seq data set, clusters were generated as described for the region enrichment analysis. In addition, peaks identified in (h)MeDIP-Seq were further subdivided into '5hmC regions with low 5mC' and '5hmC regions with high 5mC'. Peaks with an abnormally high coverage of bisulphite reads were identified with a box-whisker distribution filter and extreme outliers (ratio of >threefold above median) were excluded from the analysis. The read coverage and methylation levels underlying these peaks or different genomic features were analysed using SeqMonk.

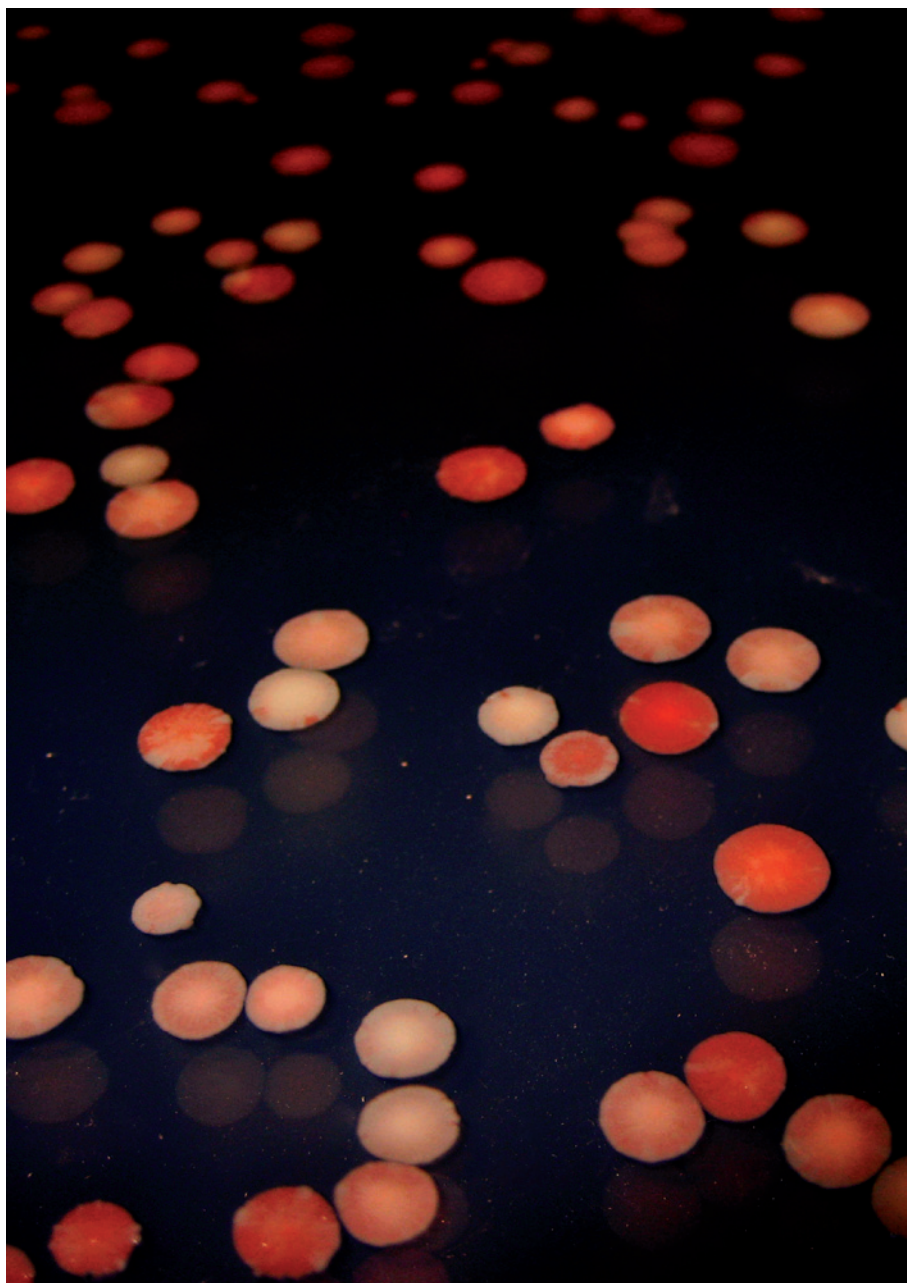
30. Tsumura, A. *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).
31. Iacovino, M. *et al.* A conserved role for Hox paralog group 4 in regulation of hematopoietic progenitors. *Stem Cells Dev.* **18**, 783–792 (2009).
32. Ting, D. T. *et al.* Inducible transgene expression in mouse stem cells. *Methods Mol. Med.* **105**, 23–46 (2005).
33. Santos, F. *et al.* Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. *Curr. Biol.* **13**, 1116–1121 (2003).
34. Ng, R. K. *et al.* Epigenetic restriction of embryonic cell lineage fate by methylation of *Elf5*. *Nature Cell Biol.* **10**, 1280–1290 (2008).
35. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
36. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$. *Methods* **25**, 402–408 (2001).
37. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

TECHNOLOGY FEATURE

THE NEXT STEP FOR THE SYNTHETIC GENOME

Biologists have copied an existing genetic code, but haven't yet commercialized it or written their own. What will it take for a tour de force to reach industrial force?

H. WANG/CHURCH LAB/HARVARD UNIV.



BY MONYA BAKER

A year ago this week, headlines trumpeted that humans had created artificial life. Scientists at the J. Craig Venter Institute in Rockville, Maryland, had chemically synthesized DNA and placed it inside a bacterial cell emptied of its own genetic material. Tests a few days after the insertion showed that the 1-million-base-pair-long synthetic genome was able to run the cellular machinery¹.

Whole-genome engineering could one day create cells unbound by biochemistry as we know it, says George Church, a geneticist at Harvard Medical School in Boston, Massachusetts. Researchers might even be able to design a new genetic code, one that could incorporate more than the 20 or so amino acids used by natural living systems. That achievement is “going to be more than an increment”, says Church, “that’s going to be a game-changer”. But current reality is more prosaic. As Venter Institute staff celebrated their cell’s first birthday with a chocolate-and-spice layer cake topped by a miniature microscope made of sugar, they were well aware that the era of synthetic genomes still faces plenty of growing pains.

Breathless headlines notwithstanding, the Venter Institute team did not create life so much as copy an existing plan. In this case, they acted more like scribes than authors. Synthetic biologists are also working on changing DNA sequences — trying to engineer microbes for practical applications such as decontaminating toxic waste, tracking down tumours or secreting biofuels — but few work with more than ten genes at a time. The story of the field so far is, “can write DNA, nothing to say”, says Drew Endy, a synthetic biologist at Stanford University in California. “We can compile megabases of DNA, but no one is designing beyond the kilobase scale.”

“Most of us are still working on a small scale because there are interesting questions there and because that’s what we have the technology to build,” says James Collins, a biomedical engineer at Boston University in Massachusetts. ►

Strains of *Escherichia coli* have been developed to produce lycopene, an antioxidant found in tomatoes.

► “We frankly don’t understand biology well enough to start designing genomes *de novo*.”

Many technologies must fall into place before researchers will be able to routinely work with even tens of genes at a time. Putting together huge DNA molecules is time-consuming and expensive, and designing biological components to perform a particular task is a challenge for parts of genes, let alone whole genomes. Transplanting DNA molecules into cells is not easy, nor is getting the DNA to ‘boot up’ once it is in place. And because the genomes will be far from perfect, researchers will need ways to tweak and test many variants.

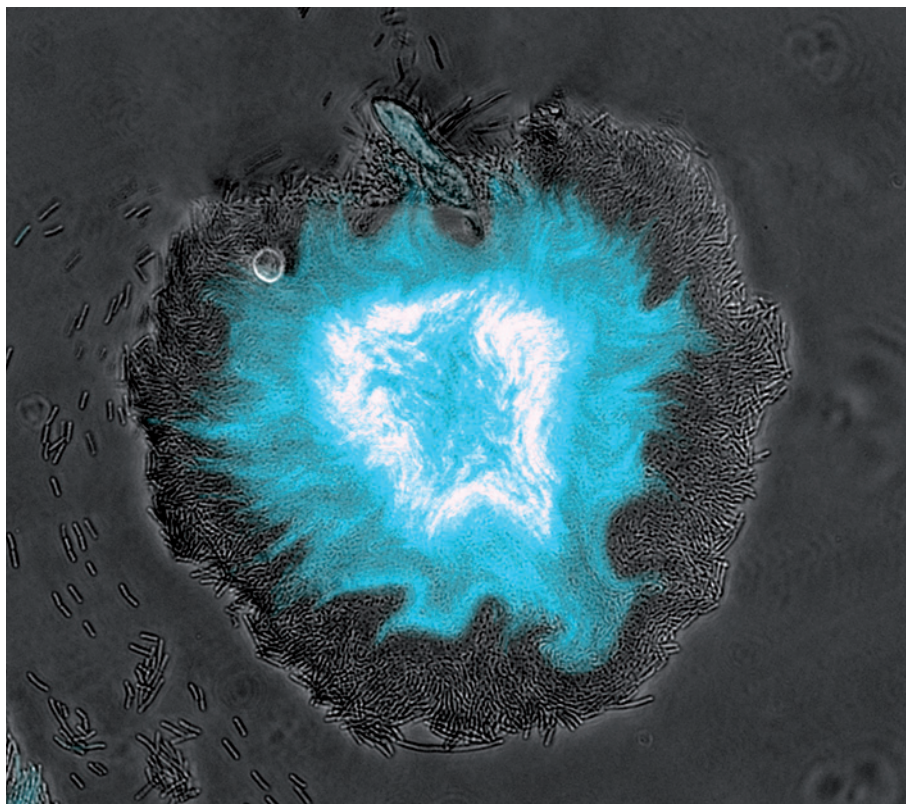
No one expects fast results, and much of the work will be tedious. The Venter Institute spent 15 years and US\$40 million creating the technology to build and transplant a genome. The 2010 paper lists two dozen authors. “This was a debugging process from the beginning,” says Craig Venter, founder of the institute. “99% of our experiments failed.” And failed experiments were costly: a single error in a million base pairs set the project back months. Not counting scientists and their equipment, four species were involved in the genome transplant: *Mycoplasma mycoides* to provide the source code, *Escherichia coli* to copy DNA pieces, baker’s yeast (*Saccharomyces cerevisiae*) to assemble them into a million-base-pair circle and *Mycoplasma capricolum* to provide the recipient shell. No wonder more synthetic biologists are thinking about parts of genes than are dreaming of constructing whole genomes.

LEARNING TO WRITE GENOMES

Synthetic biology often adopts the language of engineers: rather than talking about genes, networks and biosynthetic pathways, practitioners prefer to talk about parts, devices and modules. ‘Parts’ refer to the protein-coding section of a gene and sundry regulatory sequences that tune gene expression. A ‘device’ is an assembly of parts that together perform a particular function, often turning a protein’s production on or off. And a ‘module’ or pathway is a collection of devices that carry out more-complex functions, such as coordinating a chemical synthesis or shunting cells between ‘growth’ and ‘production’ modes.

Jeff Hasty, a bioengineer at the University of California, San Diego, used three genes to make bacteria light up in sync². Each gene is indirectly activated by the same small molecule: one controls the production of the molecule, another directs its degradation and the third makes the fluorescent protein that causes the cell to flash. The molecule diffuses between cells and coordinates bursts of protein production.

Another example of bioengineering involves a dozen or so genes from multiple species. Jay Keasling, a biologist at the University of California, Berkeley, engineered *E. coli* and yeast cells to make a precursor of the malaria drug artemisinin at one-tenth of the cost incurred



Engineered bacteria produce flashes of fluorescence, controlled by three genes in a synchronized circuit.

by the conventional method of production: extracting the natural product from sweet wormwood³. (More importantly, microbes grow faster than the plants, which are in limited supply.) Sanofi-aventis in Paris and the Institute of One World Health in South San Francisco, California, plan to start distributing the synthetic form of artemisinin next year.

But Keasling’s achievement is an object lesson in the time and expense involved. He and his colleagues began work a decade ago and had a \$43-million grant from the Bill & Melinda Gates Foundation in Seattle, Washington, to Keasling’s lab and to Amyris Biotechnologies in Emeryville, California, a company that Keasling co-founded in 2003. Researchers had to track down a previously unidentified enzyme, and engineered a dozen further yeast enzymes not just to work in *E. coli*, but also to operate at the right levels to move chemical intermediates towards a desired product without poisoning the cell or wasting resources.

To speed up such projects, the Massachusetts Institute of Technology (MIT) in Cambridge maintains a Registry of Standard Biological Parts (<http://partsregistry.org>) that lists thousands of components. However, descriptions of these parts are often incomplete, and they don’t all work as described.

To address this issue, Endy and bioengineers from the University of California, Berkeley, launched the International Open Facility Advancing Biotechnology (BioFab) in Emeryville in 2009, with a grant from the National Science Foundation. The BioFab aims to boost the

supply of working parts, both by optimizing the parts themselves and by developing systems to swiftly design genetic constructs. The goal, says Endy, is to create a set of genetic regulatory elements to precisely control the rates and levels of protein production. The BioFab currently provides 350 promoters, grouped into ten levels of protein production. Having a range of options is important, says Endy, because using the same sequences multiple times makes genetic constructs unstable. The team is assessing how these elements behave in systems and under different *E. coli* growth conditions.

Eventually, the researchers hope to create vast libraries combining variants of different parts. This will let them compare the parts’ performances, and pick the best ones. Computer analysis will then be used to model how different sequences affect gene expression, which can in turn predict how new combinations of parts will function. But the *in silico* design process can go only so far. “Models are not yet as predictive as they could be,” says Adam Arkin, a bioengineer at Berkeley and co-director of the BioFab. “In almost all cases of real application we are faced with some tinkering,” he adds. And the more parts are combined, the more unpredictable are the results.

SOME ASSEMBLY REQUIRED

The biological parts are generally easy to come by — short stretches of DNA can be ordered from a variety of companies (see ‘Making DNA on the cheap’) — but physically assembling multiple parts can be cumbersome and

Making DNA on the cheap

The description of the first chemical synthesis of a gene took up an entire issue of the *Journal of Molecular Biology* in December 1972. Then, copying an oligonucleotide of just 20 base pairs took a chemist two years¹¹. Now, researchers can order oligonucleotides 50–200 base pairs long from several vendors, for shipment within 24 hours.

"We ship tens of thousands of oligonucleotides every single day. That didn't happen 20 years ago," says John Havens, vice-president of business development at Integrated DNA Technologies in Coralville, Iowa, which supplied the oligonucleotides used by the J. Craig Venter Institute in Rockville, Maryland, to assemble the mouse mitochondrial genome⁵.

For an extra cost, oligonucleotides can be assembled into synthetic genes up to 1,500 base pairs long, which are stitched into plasmids — circles of DNA that can be propagated in bacteria. Companies including Integrated DNA Technologies, GeneArt in Regensburg, Germany, OriGene in Rockville, Maryland, and DNA2.0 in Menlo Park, California, can make genes, check for errors and ship them to clients in two weeks or less.

Prices have fallen, but much more slowly than those of multiplex gene sequencing (see 'Reading and writing DNA'). Ten years ago, synthesis cost US\$25 per base and sequencing cost \$0.25, according to data compiled by Rob Carlson, a principal at Bidesic, a consulting firm in Seattle, Washington. Last August, the figures were \$0.35 per base for synthesis and \$0.00000317 for sequencing. Synthesis at prices comparable to sequencing would greatly accelerate synthetic biology, says Christopher Anderson, a genetic engineer at the University of California, Berkeley.

Cheaper DNA should be coming.

Oligonucleotides have conventionally been synthesized individually on glass-column supports, but now microarrays allow thousands of molecules to be assembled side-by-side; microfluidics, microelectrodes or even tiny beams of light can be programmed to direct the desired chemical reaction to each spot on the microarray. LC Sciences in Houston, Texas, MYcroarray in Ann Arbor, Michigan, and CustomArray in Mukilteo, Washington, all offer oligonucleotide libraries synthesized on microarrays; so does Agilent Technologies in Santa Clara, California, but only to a limited number of collaborators.

So far, oligonucleotides produced on microarrays have been difficult to use for gene synthesis: only very small amounts of each DNA sequence are made; thousands of sequences are mixed together and hard to separate; and the error rate is usually far higher than that of conventional techniques.

Work by George Church, a geneticist at Harvard Medical School in Boston, Massachusetts, and Jingdong Tian, a synthetic-systems biologist at Duke University in Durham, North Carolina, may help to change that. Tian designed a microarray in which groups of 60-mer oligonucleotides are synthesized in small wells, facilitating their assembly into genes. These microarrays could make many variants of the same gene, which could be quickly assessed for desirable properties, such as high levels of protein expression¹². Currently, the costs of gene synthesis can become prohibitive for such experiments.

Separately, Church and colleagues at Agilent, the Wyss Institute for Biological Engineering in Boston and Stanford University in California took another approach to assembling genes from a microarray. By carefully designing the oligonucleotides and PCR primers that target them, they could selectively amplify the oligonucleotides necessary for genes¹³. This method successfully made 40 out of an attempted 42 genes for single-chain antibodies, a particularly challenging target.

Achieving that kind of performance on a commercial scale will be difficult, but the Wyss Institute is testing the waters. Later this year, it plans to use the technique to launch a service that will construct genes for as little as \$10 for a 500-base-pair piece of DNA. "We are interested in what the scientific community can do with very cheap, though imperfect, sources of DNA," says Church. **M.B.**

expensive. DNA molecules are either designed using complementary DNA sequences or mixed in with DNA complementary to the opposing ends of the molecules that are to be joined. These are combined with enzymes that cut and join DNA. Researchers can link elements using a system called BioBricks, in which sequences are cut out of circular genetic elements called plasmids by restriction enzymes specific to a particular series of nucleotides at the start and end of the sequence. The desirable parts are then stitched together into larger plasmids by other enzymes. (New England BioLabs in Ipswich, Massachusetts, sells a kit of the necessary enzymes and buffers.) Assembled sequences can then be replicated in bacteria.

Each assembled DNA piece starts and ends with the same sequences as the component parts, theoretically allowing larger and larger components to be assembled sequentially. But only three elements can be put together in a single reaction, which generally takes a couple of days. Reactions are also less successful with longer molecules, discouraging long assemblies.

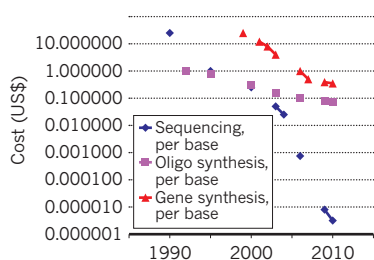
Tom Knight, a computer scientist and co-founder of start-up company Gingko BioWorks in Boston, invented BioBricks while working at MIT and has redesigned the system for industrial applications. The proprietary version can assemble up to ten parts in a single reaction, says company co-founder Barry Canton. This allows researchers to work on DNA molecules with as many as 100,000 base pairs, although most of the pathways that Gingko is working on are half that size. Just as importantly, most assembly steps can be performed by liquid-handling robots. For example, rather than bands of DNA being isolated from a gel, as in most methods, DNA molecules are collected onto and separated by suspended magnetic beads. Such automation speeds assembly and frees up lab technicians for more complicated tasks.

But BioBricks-type methods are limited by their use of restriction enzymes. Because the enzymes cut DNA whenever they encounter a particular series of nucleotides, there are 'forbidden sequences' that must be excluded from the genetic construct to avoid errant cutting. The larger a construct becomes, the harder it is to avoid such sequences. To circumvent this problem, researchers have developed assembly 'overlap' methods, in which opposite ends of molecules are joined as DNA is copied. Dozens of separate pieces of DNA can be assembled in the same reaction, often totalling a few thousand nucleotides. These methods have their own drawbacks, however. Most copy DNA using the polymerase chain reaction (PCR), which can introduce errors.

There is a bewildering array of overlap assembly techniques. 'Gibson assembly', invented by Daniel Gibson and his colleagues at the Venter Institute, allows many sequences to

READING AND WRITING DNA

The price of sequencing has fallen faster than that of making genes and oligonucleotides.



be assembled in parallel, and can even stitch together entire genomes⁴. In one demonstration, the team started with six hundred '60-mers' (oligonucleotides 60 base pairs long), and went on to assemble the 16.3-kilobase mouse mitochondrial genome⁵.

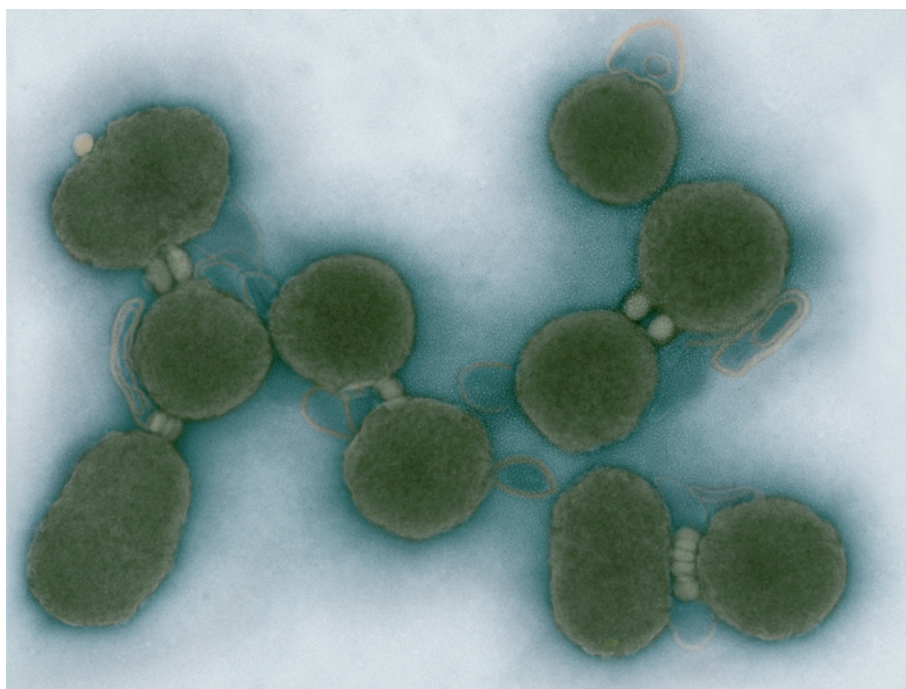
Other methods include Golden Gate Shuffling, sequence- and ligation-independent cloning (SLIC), splicing by overlapping extension (SOEing), enzymatic inverse PCR (EIPCR), overlap extension and more⁶. Some commercial kits are available: In-Fusion, from Clontech in Mountain View, California, has a mix of enzymes that can assemble 15-base-pair overlaps of any desired sequence. Life Technologies in Carlsbad, California, sells a plasmid-construction kit, MultiSite Gateway, that can join molecules with specific overlap sequences; it also markets the GeneArt High-Order Genetic Assembly System, which can assemble 10 DNA molecules, totalling up to 110 kilobases.

Researchers also design their own assembly reactions. To help this, the Joint BioEnergy Institute in Berkeley has invented a design tool, dubbed j5, that let researchers work with several DNA assembly protocols. It determines which overlap sequences to use, recommends the sequences to order from vendors and can instruct liquid-handling robots. Synthetic Genomics in La Jolla, California, which was co-founded by Venter, plans to start offering fee-for-assembly services later this year.

Assemblies larger than about 100 kilobases may be best put together inside cells, because big DNA molecules are fragile and difficult to manipulate. *In vitro* replication is also less accurate than cells' machinery. The Venter Institute team managed to assemble a 583-kilobase genome *in vitro*⁷, but it ultimately developed an *in vivo* assembly system for its synthetic genome.

Larger genomes than that of *M. mycoides* have been assembled inside cells, albeit not from synthetic starting points. In 2005, Mitsuhiro Itaya, a biochemist now at Keio University in Tsuroka, Japan, and his colleagues constructed a 3,500-kilobase cyanobacterium genome⁸. They cut the genome of the bacterium *Synechocystis* PCC6803 into large chunks and propagated them in specially prepared plasmids in *E. coli*. The plasmids were then transferred into a third species, *Bacillus subtilis*, where the DNA was stitched together.

Assembly methods aren't interchangeable. Overlap sequences that work for one method



Bacterial cells carrying a synthetic genome can grow and divide like normal cells.

often don't work for others, so researchers who run into problems with one technique have to start from scratch, says Tom Ellis, a synthetic biologist at Imperial College London. Ellis is working with Geoff Baldwin, a biochemist also at Imperial, and other colleagues to develop rules to find out which sequences will work with multiple overlap techniques, including recombination in yeast and *Bacillus*. That way, if one technique doesn't work, researchers can try others quickly.

These standards will also allow researchers to assemble DNA pieces in any order, says Ellis. Although a dictated order of assembly is fine for copying an existing genome, it does not let synthetic biologists test multiple possibilities. That issue is going to become more important as researchers move from working with thousands of base pairs to tens of thousands (see 'Sizing up synthetic DNA'). If researchers start building genomes or even large parts of genomes, they will have to think about how the DNA will wrap up on itself, and how they can place genes in chromosomes so that they end up in the right places, says Ellis. "It's a whole other aspect we'll have to uncover if we're going to do genome engineering."

EDITING IS ESSENTIAL

Jef Boeke, a molecular biologist at Johns Hopkins Medical Institute in Baltimore, Maryland, believes that genome-scale engineering is coming more quickly than many think. He is building artificial yeast chromosomes, each about the same size as the *M. mycoides* genome. Although he hasn't yet been able to design an entire new genome, he has developed techniques to make systematic alterations in existing genetic codes. "It opens the door to a lot

of imaginative change at the genome scale that wasn't possible before," he says. For example, one systematic study in 2008 deleted introns (regions within genes that don't code for protein) from many yeast genes individually, and found that the procedure had surprisingly little effect on the growth and fitness of cells⁹. Boeke wants to use his techniques to find out what will happen if all introns are removed from the genome at once.

But new possibilities introduce new problems. For the next few years, large genome assemblies are going to take months to build. With every assembly, researchers will detect unanticipated errors or realize after the fact that another sequence should work better, predicts Ellis. Then they will need to decide whether to assemble the whole genome again, or just edit it. "There has not been widespread acknowledgement in the synthetic-biology community that this is going to be an issue as we go into bigger assemblies," he says. The problem has already made itself felt: a quotation that the Venter Institute had incorporated into its synthetic genome turned out to contain a mistake, and is going to be altered.

Another use of editing is to produce and compare many gene variants. In a colourful demonstration in 2009, Church and his colleagues described a high-throughput editing system. Multiplex-automated genome engineering (MAGE) mixes bacteria with synthesized stretches of DNA that are designed to target many areas in the genome; carefully timed jolts of electricity cause the bacteria to take up the DNA as they grow in culture. Church used MAGE to alter 24 genes in *E. coli* at once, focusing on those involved in making lycopene, an antioxidant and pigment found



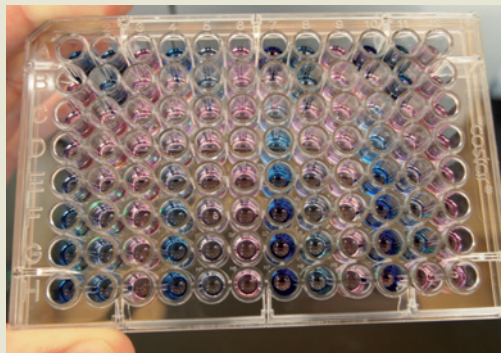
"It's quite likely that transplantation will be the unique step for each species."

Craig Venter

The useful genome

In the short term, advances in genome assembly are not likely to matter much commercially, says Rob Carlson, a principal at the engineering, consulting and design company Biodesic in Seattle, Washington. Getting the genes right is only part of the challenge; finding the best conditions for growing microbes and extracting their products is just as important. "People will build things that are as complex as can be understood and commercialized," says Carlson. Amyris Biotechnologies of Emeryville, California, is making biofuels by working with about a dozen genes, not a million-base-pair genome. "It's true that you can build something that size; it's not true that you would do that in a circumstance where you have to worry about the bottom line," says Carlson. "The problem is in getting stuff to work with the design and tinkering, rather than the assembly."

Amyris, Easel Biotechnologies in Costa Mesa, California, Gevo in Englewood,



Genetic assemblies are screened for desired effects.

Colorado, Joule Unlimited in Cambridge, Massachusetts, and LS9 in South San Francisco, California, are making biofuels and petroleum extracts from yeast, blue-green algae, *Escherichia coli* and other microbes. These efforts are backed by high-profile scientists and venture capitalists. Synthetic Genomics in La Jolla, California, has a US\$300-million partnership with

ExxonMobil and plans to use engineered microbes to make clean water, fuel and vaccines. Amyris, which went public last year, is now valued at more than \$1 billion. One company that folded, Codon Devices of Cambridge, Massachusetts, had focused not on making a product but on supplying services: providing synthetic genes to order and helping companies to develop synthetic-biology applications.

Ginkgo BioWorks in Boston, Massachusetts, originally planned to offer services for large DNA assemblies, then decided to focus on engineering the microbes themselves. Business partners and investors are less interested in new ways to assemble DNA than in better ways to manufacture products, says Barry Canton, the company's co-founder. "Companies may be shifting from a chemical synthesis platform to a biosynthesis platform," he adds. Those who control the microbes control the means of production. **M.B.**

in tomatoes. Within three days, some bacterial cells were making five times more of the red stuff than cells in the starting population¹⁰. The need for custom equipment and the difficulty of purifying transformed cells has kept researchers from widely adopting the technique, but the sheer number of genetic possibilities that can be tested using MAGE is a huge advantage, says Church. As many as 4 billion *E. coli* genomes were produced in the course of one experiment. "You're not resting on the outcome of one construct," he adds.

Mutagenesis and directed evolution of existing genomes could also help synthetic biologists to make up for current gaps in knowledge, says Collins. As more genes are brought into the system, he says, "uncertainty goes up exponentially, and you run up against the limits of what you can do modelling-wise". And although computational approaches are not yet sophisticated enough to design new genomes, they are

good at modelling existing ones, he says. This understanding could help researchers to co-opt existing cellular networks to perform desirable tasks. "We are starting to see labs recognize that there is a lot to be exploited inside the cell," says Collins (see 'The useful genome').

BIOLOGY MATTERS

The most difficult problem may well be one of the least discussed: putting the genome to work. Although Itaya has synthesized large genomes inside cells, the introduced genomes do not go on to produce proteins. Venter's group had originally chosen *Mycoplasma genitalium* for the synthesis project because its genome was, at the time, the smallest known: only 583 kilobases. But *M. genitalium* grows so slowly that the team switched to its faster-growing cousin, even though its genome is twice the size. Making the DNA is not the rate-limiting step, says Venter. "It's much more dealing with the complexity of biology versus the chemical synthesis," he says.

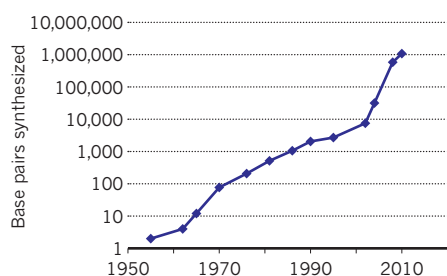
In fact, Venter thinks that adapting genomes to work in different cell types may be one of the most difficult tasks. The creation of the first synthetic cell is illustrative: the team had to remove certain enzymes from recipient cells to keep them from cutting up the foreign DNA. And moving to other species is going to be even more difficult. Unlike *Mycoplasma*, many microbes contain tough cell walls that resist the introduction of DNA. "It's quite likely that transplantation will be the unique step for each species," says Venter.

Like a child learning to write, researchers must be able to copy natural genomes before they can create new ones. One day, geneticists will be able to design code on large scales, fueling as-yet-undreamed-of applications, says Venter. "After we sequenced the genome, analysts were arguing that there was no more need for sequencing, and I argued that this was the starting point." The question of whether whole-genome synthesis will be useful will prove foolish in time, Venter believes. "It's like asking, 'why would you want to invent an airplane when people already have horses?'" ■

Monya Baker is technology editor for *Nature* and *Nature Methods*.

SIZING UP SYNTHETIC DNA

Artificial DNA has grown from two-nucleotide molecules in 1955 to more than one million in 2010.



- Gibson, D. G. *et al. Science* **329**, 52–56 (2010).
- Danino, T., Mondragón-Palomino, O., Tsimring, L. & Hasty, J. *Nature* **463**, 326–330 (2010).
- Ro, D.-K. *et al. Nature* **440**, 940–943 (2006).
- Gibson, D. G. *et al. Nature Meth.* **6**, 343–345 (2009).
- Gibson, D. G., Smith, H. O., Hutchison, C. A., Venter, J. C. & Merryman, C. *Nature Meth.* **7**, 901–903 (2010).
- Ellis, T., Adie, T. & Baldwin, G. S. *Integr. Biol.* **3**, 109–118 (2011).
- Gibson, D. G. *et al. Proc. Natl Acad. Sci. USA* **105**, 20404–20409 (2008).
- Itaya, M., Tsuge, K., Koizumi, M. & Fujita, K. *Proc. Natl Acad. Sci. USA* **102**, 15971–15976 (2005).
- Parenteau, J. *et al. Mol. Biol. Cell* **19**, 1932–1941 (2008).
- Wang, H. H. *et al. Nature* **460**, 894–898 (2009).
- Caruthers, M. H. *Biochem. Soc. Trans.* **39**, 575–580 (2011).
- Quan, J. *et al. Nature Biotechnol.* doi:10.1038/nbt.1847 (2011).
- Kosuri, S. *et al. Nature Biotechnol.* **28**, 1295–1299 (2010).

CAREERS

TURNING POINT Chemist's precision leads to new brain-function metrics **p.411**

ASK THE EXPERT Get incisive answers to career concerns go.nature.com/uuuigh

NATUREJOBS For the latest career listings and advice www.naturejobs.com



Forensic scientists must handle evidence, appear in court and understand the legal process.

FORENSICS

The call of the crime lab

Forensic scientists can work in academia, government and the private sector, but the field is competitive.

BY VIRGINIA GEWIN

Cedric Neumann has witnessed first-hand his field's scientific coming of age. "When I started my undergraduate degree, nobody wanted to work in forensic science; there was only a handful of programmes in the world," he says. Instead, "police officers were trained to work in a lab". But in the late 1990s, around the time that Neumann began his graduate studies at the University of Lausanne in Switzerland, which has the oldest forensic-science programme in the world, dozens of undergraduate and postgraduate programmes worldwide began to churn out forensic scientists.

Two things have made forensics a more visible — and fashionable — career choice, says Neumann. An increased focus on using DNA

technologies to solve crimes has sparked a demand for properly trained biologists. And several television drama series that glorify forensic science have generated so much interest in the field in the past ten years that students have been flocking to study it.

Forensic science encompasses a range of disciplines — including DNA analysis, examination of fingerprints or footwear impressions, and firearm analysis — used to solve criminal cases. In recent years the field has come under scrutiny amid calls for peer review to establish the reliability and accuracy of many forensic methods — and to develop them further.

The discipline has been subject to market pressures. While Neumann was doing his PhD in his spare time, he also worked at Britain's

Forensic Science Service (FSS), a government-owned company that provided the bulk of forensics services and research in the country. But by last year the service was losing £2 million (US\$3.3 million) a month, and was deemed no longer cost-competitive against private-sector rivals who typically analyse only a few types of forensic evidence. The FSS is slated to close next March, making England and Wales the only parts of the world with a totally privatized forensic-science market.

Neumann, now developing statistical tools to identify fingerprints at Pennsylvania State University in University Park, is among the many forensic scientists who are concerned that privatization could compromise the quality of the science that has been achieved by services such as the FSS. For example, Neumann says, if the focus turns to only highly commercial products, such as DNA profiles, there may be less emphasis on analysing sources of evidence that require more specialized training, such as handwriting examination or tool marks, thus reducing the availability of the broad knowledge often needed to solve crimes.

Still, crime isn't going anywhere, and neither is the need for forensic scientists — if anything, demand for their skills has increased over time. Today, forensic scientists can find jobs in government labs, private industry and, increasingly, academia. Yet a surplus of trainees, an economic downturn and ever-shifting political agendas make career prospects in this field difficult to predict. The mixture of scientific acumen and forensic training needed to carve out a successful sleuthing career will depend on an applicant's personal goals.

UNSTEADY DEMAND

Although the field is growing worldwide, that growth is uneven. "The demand for forensic scientists is good, but spotty," admits Jay Siegel, director of the forensic and investigative sciences programme at Indiana University–Purdue University Indianapolis.

For example, the UK forensics scene is in the middle of a mighty shake-up as it prepares for the FSS's closure — a move that will result in 1,600 lay-offs. The turnover promises to open opportunities for new graduates as private companies recruit to fill the gap. "We anticipate that many people working at the FSS may choose not to relocate to other providers," says Steve Allen, managing director of LGC Forensics, a private forensic-services company headquartered in Teddington, UK. So far, says Gillian Tully, head of research ►

► and development at the FSS, some 90% of the 300 staff already laid off at the three shuttered FSS sites in Birmingham, Chepstow and Chorley have left the profession completely; some retired and others didn't want to move or were disgruntled by the privatization, she says. Allen hopes that LGC Forensics, Britain's largest private forensics firm with 550 employees, will increase its staff by 50–100% as the FSS workload is divvied up. "We're out there right now actively recruiting at all levels," he says.

This short-term recruitment push, however, might mask other long-term trends. Public-sector budgetary constraints continue to stymie recruitment worldwide. "We had a big growth phase a few years back. Now we are in a tight fiscal climate and I expect limited recruitment in the next two to three years," says Gary Pugh, director of forensic services at New Scotland Yard in London, the headquarters of the Metropolitan Police, who will rely solely on his in-house forensics team and private providers after the demise of the FSS. In the United States, economic conditions have grown so dire that cuts are being made at many of the 400 publicly funded crime labs that support local, state and federal branches of law enforcement. Alabama has been the most severely affected state, closing three crime labs to save money in the state budget.

But if the economy is pushing recruitment down in some regions, politics and crime patterns can easily prompt hiring elsewhere. For example, New Jersey — where the governor, Chris Christie (Republican), is the former US attorney for the state, the chief federal law-enforcement officer — created 29 positions in the state Office of Forensic Sciences amid budget cuts. And, despite budget woes, administrators in Los Angeles, California, are recruiting for the last of 26 DNA technician positions

WHAT APPLICANTS NEED TO KNOW

The case for a forensics PhD

Advantages

- A PhD is a must if you know you want to do forensics research at an academic institution.
- A doctorate can make you a more credible witness in court.
- Some laboratories, especially US federal labs, tend to promote people with PhDs more quickly and more willingly than those without, says Jay Siegel, director of the forensic and investigative sciences programme at Indiana University–Purdue University Indianapolis.

Disadvantages

- There aren't many university programmes that offer a PhD in forensic science — and those that exist are all in Europe. A more typical route to employment, especially in the United States, is to get a postgraduate degree in chemistry or biology, with a specialism in forensics.
- A PhD isn't required for a typical forensics job in a crime lab — and it can be a deterrent to getting hired, because employers might be concerned that the candidate will get bored or leave for a job in academia, says Siegel. **V.G.**

intended to reduce the city's backlog of 6,000 sexual-assault DNA-collection kits that need analysis, says Greg Matheson, director of the Los Angeles Police Department Crime Lab.

Despite federal budget cuts, Vermont Senator Patrick Leahy (Democrat) introduced a federal bill in January to boost forensic research needed to strengthen the quality of evidence routinely used in the criminal justice system. But Siegel doubts that the bill will pass in this climate.

GETTING PROPER TRAINING

The 'CSI effect' — the tremendous interest in the field aroused by the US television programme *CSI: Crime Scene Investigation* and similar dramas — has spurred a flood of applications for forensic-science jobs. "If we have an opening for a forensic chemist, we can easily get 200–300 applications for that position," says Michael Medler, laboratory director of the forensic-services agency in Indianapolis, Indiana. The agency, an independent government entity, hires crime-scene specialists, forensic analysts and technicians.

As a result of the interest, employers have their pick of the talent, and increasingly choose applicants with postgraduate degrees. "In the past year, most forensic labs have become bigger, and the scientific requirements for applicants continue to become more rigorous," says Jan De Kinder, director of the Belgian National Institute of Criminalistics and Criminology in Brussels, a research body within the justice department that conducts original research and develops new forensic techniques. He says that half of the 100 people working in his lab have MSc or PhD degrees — a worldwide trend, he adds. And whereas a PhD is not a requirement for many of the jobs in a forensic lab, a lack of one can affect aspirations (see 'The case for a forensics PhD').

The scientific skills most in demand include DNA profiling and mass spectrometry for analysing the chemistry of trace specimens, as well

as statistical analysis to identify patterns found in other types of evidence, such as fingerprints. Siegel says that the biggest trend in academic forensics research at the moment is the attempt to validate the techniques used in such pattern-evidence analysis; a 2009 report by the US National Academy of Sciences highlights this need.

Universities have seized on the growing interest in forensics as a money-making opportunity. Hundreds of undergraduate and MSc forensic-science programmes now exist, but their quality varies widely. The American Academy of Forensic Sciences in Colorado Springs, Colorado — the field's professional organization — established the Forensic Education Programs Accreditation Commission (FEPAC) in 2004 to accredit undergraduate and graduate degree programmes that meet minimum standards of excellence.

Since then, 35 have been accredited, and that could grow to 50 within the next five to ten years, says Siegel. "The standards are quite rigorous — the faculty and instruments needed to teach forensic science are pretty expensive," he adds.

Some employers tend to hire traditional chemistry or biology graduates rather than graduates of forensic-science programmes. "It's really important to have a good scientific mindset and experience in biology or chemistry; we can train them in the forensic part," says Allen. Positions at private companies, often further removed from the field's judicial responsibilities, may not demand the forensics know-how required at public agencies.

Chris Hassell, director of the US Federal Bureau of Investigation laboratory in Quantico, Virginia, echoes the importance of a scientific background. "We don't generally target people coming out of forensics disciplines; we're looking for good scientists," says Hassell. He says that a candidate's scientific publication record can set her or him apart from the crowd — for example, if it includes a paper addressing

P. BOUDRY



Jan De Kinder runs a Belgian institute that develops new forensic techniques.

the validity of a specific forensic technique.

Commitment is key. “Unlike other areas of science, in forensics a person’s credibility is called into question daily in a court of law,” says Medler. In addition to mastering a range of scientific techniques, he says, forensic scientists must be able to identify the most probative pieces of evidence at the crime scene, must know how to document who has physical possession of evidence and why, have knowledge of the legal process and have the ability to communicate on a court stand. Much of that training must be acquired on the job.

In Europe, the training requirements for crime-lab analysts vary depending on which body has authority over the forensics operations. For example, in France, Italy and Spain, forensic services are provided by the police; until recently, only trained police officers could work in crime labs. However, in Belgium, forensic labs are under the purview of the justice department.

Applicants with criminal records or who fail drug tests face dim prospects. Matheson says that background checks disqualify up to two out of every ten candidates.

CLOSING THE GAP

Forensic science is considered a young field. Police labs, frequently inundated with caseloads, are often simply unable to perform much-needed research. And although there is a growing amount of forensics research in academia, interactions between practitioners and researchers can be limited.

But as the number of forensic-science programmes at universities grows, and the PhD and MSc students chip away at research needs, the field’s scientific footing is expanding. “The advantage of having more university training programmes in forensics is the increase in research activities,” says De Kinder. Unfortunately, researchers still struggle to find funding.

“To better our profession we need to do two things: encourage people with PhDs to get into forensics and overcome the disconnect between academia and the practising field,” says Larry Quarino, chair of FEPAC and director of the forensic-science programme at Cedar Crest College in Allentown, Pennsylvania. He advocates the creation of a sabbatical that would allow practising forensic scientists to conduct academic research necessary for their positions.

“For a scientific discipline to be a living discipline, it needs to conduct research,” says Pierre Margot, head of the school of criminal justice at the University of Lausanne. “As long as researchers are working on the needs of tomorrow,” says Margot, “I’m not too worried about the state of the job market today.”

Virginia Gewin is a science journalist based in Portland, Oregon.

TURNING POINT

Jill Venton

Jill Venton, an analytical chemist at the University of Virginia in Charlottesville, received the 2011 Society for Electroanalytical Chemistry Young Investigator Award in March for her efforts to develop sensors able to probe neurotransmitters in fruitflies.

As an analytical chemist, do you find neurochemistry messy?

Analytical chemists develop methods to quantify the composition and structure of matter, and I definitely think like an analytical chemist — I like precise measurements with small error bars. But life does not take place in a beaker, and I knew early in my career that I wanted to apply my skills to biology. I did my PhD in analytical chemistry with a neuroscience focus and found that I liked the field, so I followed up my degree by doing a postdoc supervised jointly by a chemist and a neuroscientist. By comparison with chemistry, neuroscience is messy. It’s more exploratory, which often doesn’t lend itself to nice, neat experiments, because we know so little about the brain — but it has been fun and challenging to use my talent for precision to help develop ways to measure brain functions.

How do you get your research ideas?

Some come from colleagues. For example, a neuroscience colleague wanted to measure neurotransmitters in the fruitfly brain and challenged me to help him find a way to do it. I had never thought of it before, but I was exploring techniques to measure fast changes in neurotransmitters in the mammalian brain, so I thought I could tackle it. Other ideas come from the need to keep pushing technology further and exploring the boundaries of what new methodology can tell us about neuroscience.

What’s your strategy for winning early-career awards?

I have applied for a lot of young-investigator awards, and certainly have not won them all. When I started out, I applied indiscriminately for any funding or award. I was lucky to get a US National Science Foundation career award early on, which helped to give my lab a foundation. Once I got that, I became pickier in terms of which awards to seek, because I didn’t have infinite amounts of time to apply to them. At the moment, I rely on national funding agencies for my bread and butter, and



apply for awards that have a certain level of prestige to supplement that.

You are awaiting a decision on tenure now. Was the tenure process what you expected?

I knew that the tenure committee would look at grants and publications, and that there would be significant emphasis on letters written on my behalf from people outside this institution. Many people do what’s called a ‘tenure tour’ in the year or so before they go up for tenure, working to raise their profiles and build a reputation in the field to ensure those positive tenure letters. I had a baby a year and a half before I went up for tenure, so my ability to travel was limited and I was more selective about where I went. For example, rather than presenting at single universities, I went to a Gordon Research Conference — an international gathering of scientists to discuss the frontiers of research. Before getting pregnant, I spent time networking by meeting people at conferences and organizing workshops or symposia.

Analytical chemistry is a male-dominated field. Does that pose challenges?

Yes. I’m one of only three women in a department of about 30 — and the only woman with a child. But it is very typical in chemistry for women to hold only 10% of the academic positions. Still, this department has accommodated my efforts to set a flexible schedule to balance work and life. The biggest challenge is that there weren’t — and still aren’t — many role models, successful female researchers. I had to look to biology and neuroscience for those. ■

INTERVIEW BY VIRGINIA GEWIN

SURVEILLANCE

The word on the street.

BY JULIAN TANG

“The problem with running a country of 80 million people is that it’s difficult to know what people are thinking — I mean, really thinking,” said the Prime Minister, thoughtfully.

Henry Irvin cleared his throat. “If I may make a suggestion, Sir?” he started, smoothing his tie and sitting up a little straighter. “We know that what people say in public, particularly when asked to express their views specifically, may not represent what they really think or feel. This is not necessarily a deliberate intent to lie or mislead, but, more often than not, it is an attempt to comply with their current peer-group beliefs or teachings — like being with your friends at school or your colleagues at work. Except for a few outstanding individuals, this seems to be the norm.”

The PM listened intently.

“However,” Henry continued, “when they believe that they are really anonymous, such as in Internet chat rooms, their real beliefs are often expressed — particularly in response to key questions. These might be about anything from the current state of the economy, their favourite football team, their friends, colleagues, et cetera, et cetera...”

“So what are you proposing, Henry?” asked the PM, cautiously.

“Well, if you really want to know what the people are thinking, Sir, you could set up your own Internet chat rooms to encourage individuals to express themselves, anonymously, about various topics of specific interest to you and monitor their responses to key questions.”

The PM sighed, disappointed. “This is nothing new Henry — this has been tried before and nothing really serious is ever discussed by serious people in these chat rooms.”

Henry paused for a few seconds before replying. “I’ve been talking to some people at GCHQ and they have an interesting idea. You’ve heard of genomics, proteomics and metabolomics, right, Sir?”

The PM nodded.

“Well, the smart guys there have come up with a new speciality, ‘grammaromics’, consisting of verbomics, nounomics, adjectiveomics and other subspecialties.”

“Are you pulling my leg, Henry?” asked the PM, only half-jokingly.

Henry shook his head. “Not at all, Sir. In fact, they’ve tried some of their algorithms



on the text from some of these Internet chat rooms already. They’ve found that as long as these individuals stay online for a while and type a minimum number of diverse words and responses, they can reliably recognize any particular individual by the way they use their English constructs. Of course, they cannot identify the individuals themselves by this method alone, but...”

Henry paused again, for he knew that what he was about to say would not make the PM particularly happy.

“... well, they hacked into a limited number of e-mail servers to see if they could identify these chat room users by matching these ‘grammaromes’ to any particular e-mail text — just like trying to match fingerprints or DNA sequences. They then back-traced these individuals’ IP addresses from their chat rooms and their e-mails to see if they were one and the same individual.”

“You mean that GCHQ invaded their privacy to test a hypothesis?” the PM demanded, severely. Then, more curiously, he asked: “Did it work?”

Henry breathed a sigh of relief. “Yes Sir. Remarkably well, in fact. It seems that individuals develop their own unique way of expressing themselves in writing that remains more or less unchanged for life — just like their fingerprints or DNA.”

“That’s amazing!” exclaimed the PM, genuinely surprised. “If we know that people express themselves more truly in Internet chat rooms and on e-mail, when they think they’re anonymous, and if we can monitor these ‘genuine’ communications then ... Henry, if this works, you truly are a genius!”

Over the next few

years of his first term, the PM’s popularity soared. His policies were adopted in record time and it seemed that he could do nothing wrong. Even the typically cynical British media were lost for words.

Then one morning, there was a soft knock on his office door. “Come in Henry!” he yelled cheerfully, waving his pen.

Henry entered, looking concerned. “Sir,” he began without ceremony. “My colleagues at GCHQ have noticed some unusual patterns of Internet chat room and corresponding e-mail activity.”

The PM motioned Henry to the seat in front of him. “What is it, Henry?” he asked, now serious.

“We think that someone, possibly via a leak from either GCHQ or this office, may have got wind of our clandestine Internet public-opinion surveillance strategy.”

“Why would they think that, Henry?” retorted the PM, impatiently. “How would anyone outside our inner circle know or suspect anything? There’s been no publication in that popular science journal — what’s its name — *Nature*, yet has there?”

“Not as far as I know, but that doesn’t mean that no one has figured this out by themselves. After all, you are doing remarkably well in the opinion polls, Sir — just about the most popular PM in British history.”

“So, what evidence do they have to make them think that anyone out there is trying to manipulate this Internet surveillance system?”

“Well, it seems that an online consensus is building that we should consider abolishing VAT and income tax, as well as providing a Jaguar or Aston Martin to everyone passing their driving test, starting from the next tax year...”

The PM digested this information for a moment then sat back in his chair, chuckling. “I guess it was too good to last, eh, Henry?”

Henry allowed himself a rare smile. “So, back to business as usual then, Sir?”

“You read my mind, Henry,” said the PM, still chuckling and picking up his pen again as Henry closed the office door quietly behind him. ■

Julian Tang is a clinical/academic virologist who has had several stories in *Futures*. Some that didn’t make the final cut can be found in an anthology, soon to be available at amazon.com.

➔ **NATURE.COM**
Follow *Futures* on
Facebook at:
go.nature.com/mtoodm

Neural crest regulates myogenesis through the transient activation of NOTCH

Anne C. Rios¹, Olivier Serralbo^{1*}, David Salgado^{1*} & Christophe Marcelle¹

How dynamic signalling and extensive tissue rearrangements interact to generate complex patterns and shapes during embryogenesis is poorly understood^{1–3}. Here we characterize the signalling events taking place during early morphogenesis of chick skeletal muscles. We show that muscle progenitors present in somites require the transient activation of NOTCH signalling to undergo terminal differentiation. The NOTCH ligand *Delta1* is expressed in a mosaic pattern in neural crest cells that migrate past the somites. Gain and loss of *Delta1* function in neural crest modifies NOTCH signalling in somites, which results in delayed or premature myogenesis. Our results indicate that the neural crest regulates early muscle formation by a unique mechanism that relies on the migration of *Delta1*-expressing neural crest cells to trigger the transient activation of NOTCH signalling in selected muscle progenitors. This dynamic signalling guarantees a balanced and progressive differentiation of the muscle progenitor pool.

Early skeletal muscle (the primary myotome, which is composed of mononucleated post-mitotic muscle fibres, or myocytes) is formed from the generation of muscle cells at the four borders of the dermomyotome, the dorsal-most epithelial compartment of somites^{4–10}. Most of the dermomyotome undergoes an epithelial to mesenchymal transition that leads to the emergence of a population of resident muscle progenitors that massively contributes to the growth of all trunk muscles^{11–14}. The medial border of the dermomyotome, the dorsomedial lip (DML), remains epithelial for a considerable period of time, during which it generates muscle cells that contribute to the growth of the primary myotome. DML stem/progenitor cells can adopt two fates during the first days of embryonic muscle development^{4–6}: to self-renew and remain in the epithelial border of the dermomyotome or to translocate in the myotome and undergo terminal myogenic differentiation. How this balance is regulated is unknown.

In the chick embryo, the epithelial DML population comprises a majority (77%) of PAX7-positive cells interspersed by (23%) PAX7/MYF5-positive cells (Supplementary Figs 1a–e, 2a–e). In the transition zone, cells shut-off the expression of PAX7, but maintain MYF5 expression. Fully elongated myocytes express skeletal muscle myosin heavy chain (MyHC; also known as MYC). NOTCH family members are expressed in the DML, the transition zone and the myotome during the first phase of myogenesis (Fig. 1a)¹⁵. The NOTCH target genes *HES1*/*cHairy2* and *lunatic fringe* (*LFNG*) are expressed in a salt-and-pepper pattern within the DML. Many transition zone cells express *HES1*, whereas *LFNG* expression is low in this region. Their expression is low in the myotome (Fig. 1a, b and Supplementary Fig. 2f). Both genes act as bona fide NOTCH targets in somites, as their messenger RNA expression is upregulated after electroporation of a constitutive form of NOTCH (NOTCH intracellular domain (NICD); Supplementary Fig. 2g–i, m–o). To quantify the distribution of NOTCH activity, we electroporated a NOTCH reporter construct consisting of the mouse *Hes1* promoter region upstream of a destabilized GFP (d2EGFP; half life, 2 h), that efficiently responds to NOTCH activation and inhibition (Supplementary Fig. 2j–l). We co-electroporated a

human histone H2B (H2B)–RFP reporter gene driven by an ubiquitous promoter to evaluate the normal distribution of electroporated cells. After 24 h, 11% of H2B–RFP-positive cells were *HES1*–d2EGFP-positive (Fig. 1c, d and Supplementary Fig. 2l). The H2B–RFP-positive cells were distributed among PAX7- (60%), MYF5- (46%) and MyHC-positive (10%) populations. In contrast, nearly all (92%) *HES1*-positive cells were MYF5-positive (distributed in the DML and the transition zone), whereas only 24% and 2% expressed PAX7 and MyHC, respectively (Fig. 1e–i and Supplementary Fig. 3a, b). We followed the morphogenic movements of NOTCH-activating cells using live video microscopy. Epithelial cells that activated the NOTCH reporter in the DML rapidly translocated in the transition zone (Fig. 1j–l and

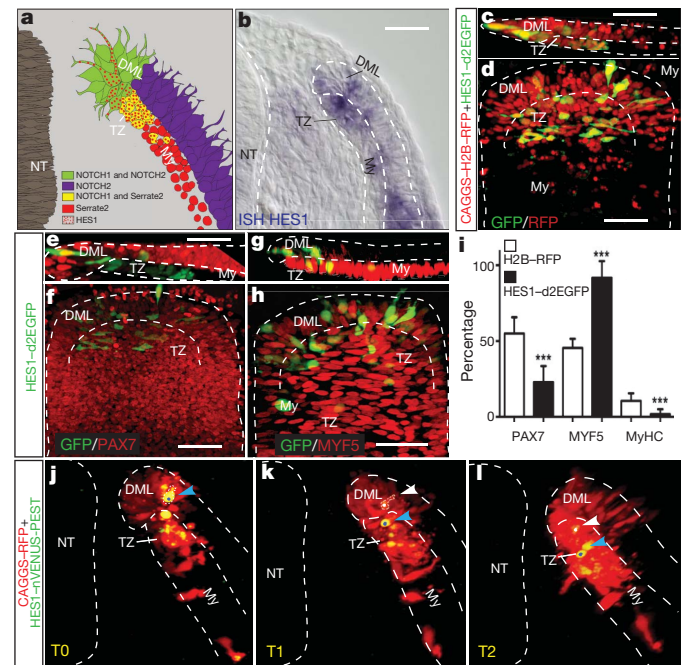


Figure 1 | Notch is active during early myogenesis. **a**, Scheme showing the expression of NOTCH signalling family members in the DML, transition zone (TZ) and myotome. *Serrate2* is also known as *JAGGED2*. **b**, Expression of chick *HES1*/*cHairy2* in the DML and the TZ. ISH, *in situ* hybridization. **c–h**, Confocal stacks showing the expression (in green) of a *HES1*–d2EGFP reporter construct and (in red) RFP (**c, d**), PAX7 (**e, f**) and MYF5 (**g, h**) in dorsal (**d, f, h**) and transverse (**c, e, g**) views of somites 24 h after electroporation. **i**, Quantification of **c–h**. Error bars show standard deviation (s.d.). *** $P < 0.0001$. **j–l**, Time-lapse confocal analysis (see Supplementary Movie 1) showing the translocation of two NOTCH-activating DML cells (blue and white arrowheads) into the transition zone. T0, start of the movie; T1, 4 h 50 min after the start; T2, 10 h after the start. My, myotome; NT, neural tube. Scale bars, 50 μ m.

¹EMBL Australia; Australian Regenerative Medicine Institute (ARMI), Monash University, Building 75, Clayton, Victoria 3800, Australia.

*These authors contributed equally to this work.

Supplementary Movie 1). We followed their fate as they further differentiated, using a construct that contains the *Hes1* promoter region upstream of a stable GFP (EGFP half life of over 24 h). After 24 and 48 h, the proportion of MyHC-positive myocytes was more than twice (24 h: 34%; 48 h: 60%) that of control RFP-electroporated embryos (24 h: 13%; 48 h: 28%; Supplementary Fig. 3c–g), further indicating that activation of NOTCH signalling is associated with myogenesis.

Altogether, this indicates that NOTCH signalling is activated in DML cells that engage in the myogenic program before they translocate into the transition zone. NOTCH signalling remains active in the transition zone and is extinguished before cells undergo terminal myogenic differentiation and elongate into myocytes.

We inhibited NOTCH activity in somites using a truncated, dominant-negative form of the NOTCH transcriptional co-activator mastermind (DN MAML1)^{16,17} and small interfering RNAs (siRNAs) against *NOTCH1* (ref. 18). DN MAML1 and the *NOTCH1* siRNA gave similar results one day later, that is, a drastic reduction of myogenic differentiation (Fig. 2). This was characterized by a sharp reduction of MYF5-positive cells (7% DN MAML1; 3% siRNA *NOTCH1*), compared to controls (49% CAGGS-IRES-EGFP; 53% siRNA luciferase), and by a halt of terminal differentiation (0% MyHC-positive cells for DN MAML1 and siRNA *NOTCH1*; controls: 12% CAGGS-IRES-EGFP; 8% siRNA luciferase; Fig. 2s, t), with no change in dermomyotome cell proliferation (Supplementary Fig. 4a–d). Virtually all cells in which

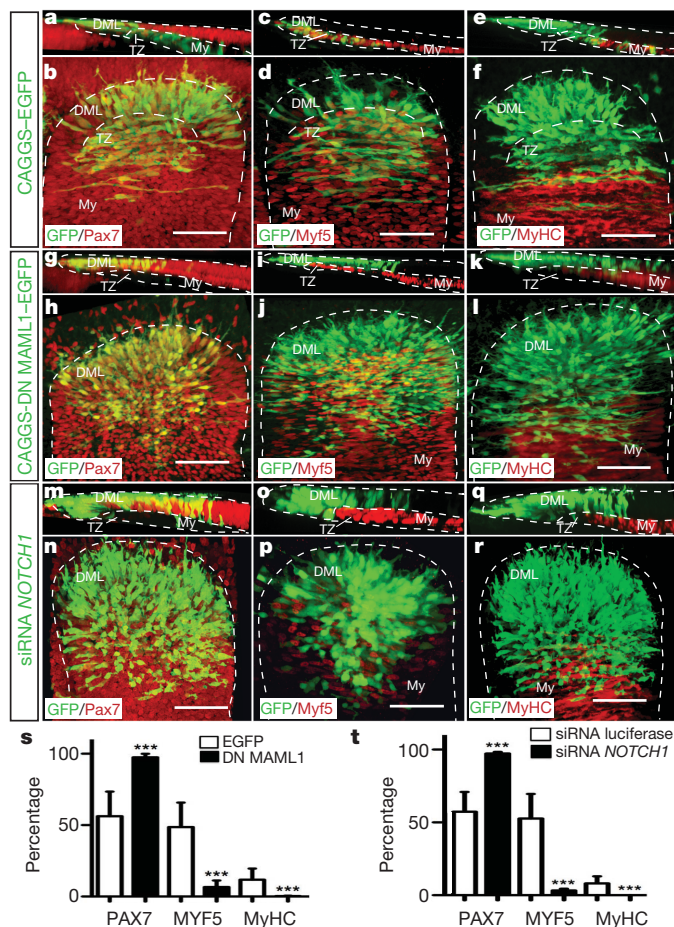


Figure 2 | NOTCH signalling is necessary for myogenesis. a–r, Confocal stacks of somites in dorsal (b, d, f, h, j, l, n, p, r) and transverse (a, c, e, g, i, k, m, o, q) view, 24 h after electroporation of (in green) CAGGS-EGFP as controls (a–f), DN MAML1 (g–l) and siRNA against chick *NOTCH1* (m–r), and stained (in red) for PAX7 (a, b, g, h, m, n), MYF5 (c, d, i, j, o, p) and MyHC (e, f, k, l, q, r). s, Quantification of a–l. t, Quantification of m–r, siRNA against luciferase as controls are not shown. Error bars show s.d. *** $P < 0.0001$. Scale bars, 50 μ m.

NOTCH signalling was inhibited remained epithelial in the dermomyotome (98% DN MAML1; 97% siRNA *NOTCH1*; controls: 56% CAGGS-IRES-EGFP; 57% siRNA luciferase). Altogether, this strongly indicates that NOTCH signalling is necessary for the initial phases of myogenesis of DML cells.

We induced a gain of NOTCH function by electroporating NICD in newly formed somites. One day later, most electroporated cells had translocated in the transition zone (Fig. 3a–f); however, most (83%) expressed the dermomyotomal marker PAX7 (Fig. 3a, b, g), and only a few (3%) were MYF5 positive (Fig. 3c, d, g). Although few electroporated cells entered the myotome region, they did not elongate and never (0%) initiated MyHC expression (Fig. 3e–g). This result is coherent with studies that showed that NOTCH signalling inhibits muscle differentiation in various contexts^{19–21}. However, characterizing the electroporated cells 6 h after electroporation of NICD, we observed a robust increase in the proportion of electroporated cells expressing MYF5 (80%; controls, 17%; Fig. 3h–k and Supplementary Fig. 5n). Strong MYF5 activation was maintained 12 h after electroporation (89%; controls, 20%; Fig. 3l–o and Supplementary Fig. 5n). After 6 h, all MYF5-positive electroporated cells were positioned in the epithelial DML (Fig. 3j), at 12 h, most electroporated cells had translocated in the transition zone (Fig. 3n). The same observations were made with MYOD (Supplementary Fig. 5a–m). Altogether, this indicates that the first steps of myogenesis (the activation of MYF5 and MYOD) are promoted by a short activation of NOTCH signalling. However, a sustained activation of NOTCH reverses the myogenic program, resulting in a downregulation of MYF5 and MYOD expression and a return to a PAX7-positive state.

To prove this, we used a doxycyclin-inducible system to drive NICD expression. In the continuous presence of doxycyclin, NICD expression was maintained in electroporated cells and, consistent with our previous observation (Fig. 3a–f), most of them translocated in the transition zone but did not maintain MYF5 expression (6%, Fig. 3r, s, v; controls, 42% MYF5-positive, Fig. 3p, q, v). When doxycyclin was removed, NICD was strongly expressed 6 h later, but was almost undetectable after overnight incubation (Supplementary Fig. 6c, d, f). Remarkably, after this transient activation of NOTCH signalling, most electroporated cells had translocated in the transition zone and the myotome and nearly all (97%) expressed MYF5 (Fig. 3t–v). In addition, electroporated cells that were positioned in the myotome had elongated into myocytes, indicating that they initiated terminal differentiation. The lack of electroporated cells in the DML (arrowheads in Fig. 3t) indicates a depletion of the DML progenitor cell population and suggests that the pulse of NOTCH signalling massively disrupted the balance between maintenance and differentiation of this cell population. This shows that NOTCH signalling displays a complex behaviour on myogenesis, acting as a potent stimulator of the myogenic program for DML cells, but only during a limited time window.

In search for a signal controlling the mosaic activation of NOTCH we observed in the DML, we noted that neural crest cells that migrate in close proximity to the DML express the NOTCH ligand Delta1 (*DLL1*) in a salt-and-pepper pattern (Fig. 4a and Supplementary Fig. 8a, b). A provocative hypothesis was thus that migrating *DLL1*-expressing neural crest cells may activate NOTCH signalling in selected progenitors within the DML. We eliminated the neural crest cell population by electroporating into the neural tube a diphtheria toxin fragment A complementary DNA (DTA) under the control of a neural-crest-specific promoter (Supplementary Fig. 7a–f). This led to a considerable decrease in the expression of MYF5 on the electroporated side (Fig. 4b, arrowheads, and Supplementary Fig. 8c; $n = 13/15$). The inhibition of non-canonical, planar cell polarity (PCP) WNT signalling in *Xenopus* affects neural crest migration²² without affecting its induction. In the dorsal neural tube, we electroporated a mutant form of the WNT intracellular effector Dishevelled that specifically inhibits the WNT/PCP pathway^{23–25}. This led to a considerable reduction in MYF5 expression compared to the control side (Fig. 4c and Supplementary

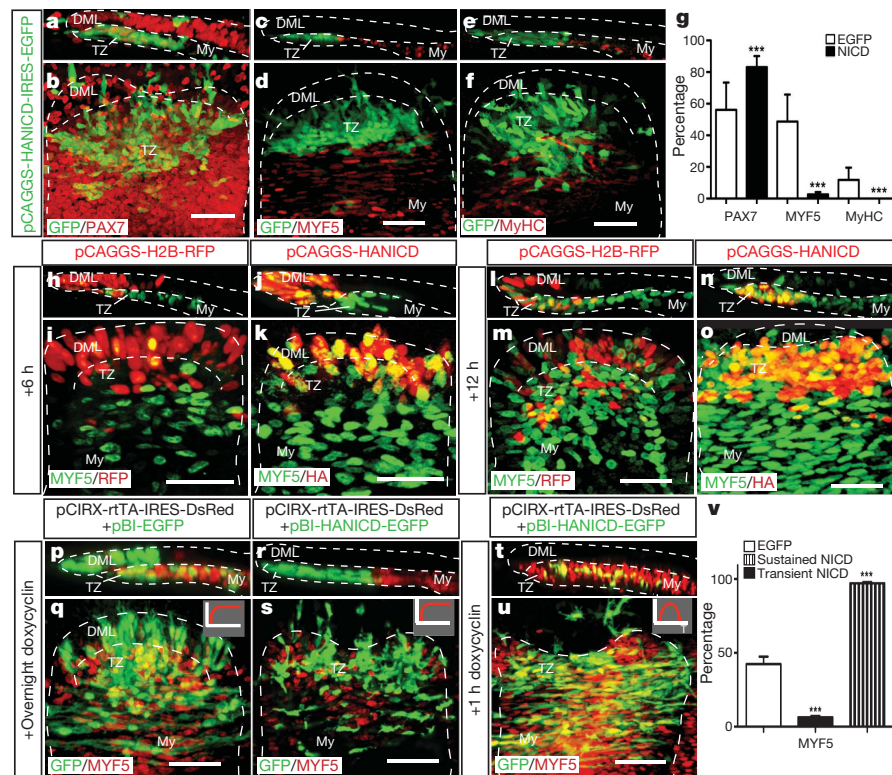


Figure 3 | Myogenesis requires the transient activation of NOTCH. **a–f**, Confocal stacks of somites in dorsal (**b**, **d**, **f**) and transverse (**a**, **c**, **e**) view 24 h after electroporation of NICD (in green), and stained (in red) for PAX7 (**a**, **b**), MYF5 (**c**, **d**) and MyHC (**e**, **f**). **g**, Quantification of **a–f**. Error bars show s.d. *** $P < 0.0001$. **h–o**, Time-course analysis of MYF5 expression (in green) in dorsal (**i**, **k**, **m**, **o**) and transverse (**j**, **l**, **n**) view, 6 h (**h–k**) and 12 h (**l–o**) after electroporation of CAGGS-H2B-RFP as control (**h**, **i**, **l**, **m**) or HA-NICD (**j**, **k**, **n**, **o**). **p–u**, Confocal stacks of somites in dorsal (**q**, **s**, **u**) and transverse (**p**, **r**, **t**) view electroporated with a doxycyclin-expression-inducible system. Embryos were electroporated with an empty vector as controls (**p**, **q**) or HA-NICD (**r**, **u**), treated for 1 h (**t**, **u**) or overnight (**p**, **s**) with doxycyclin and stained for GFP (in green) and MYF5 (in red). **v**, Quantification of **p–u**. Scale bars, 50 μ m.

(**j**, **k**, **n**, **o**). In red, staining for RFP (**h**, **i**, **l**, **m**) or HA (**j**, **k**, **n**, **o**). **p–u**, Confocal stacks of somites in dorsal (**q**, **s**, **u**) and transverse (**p**, **r**, **t**) view electroporated with a doxycyclin-expression-inducible system. Embryos were electroporated with an empty vector as controls (**p**, **q**) or HA-NICD (**r**, **u**), treated for 1 h (**t**, **u**) or overnight (**p**, **s**) with doxycyclin and stained for GFP (in green) and MYF5 (in red). **v**, Quantification of **p–u**. Scale bars, 50 μ m.

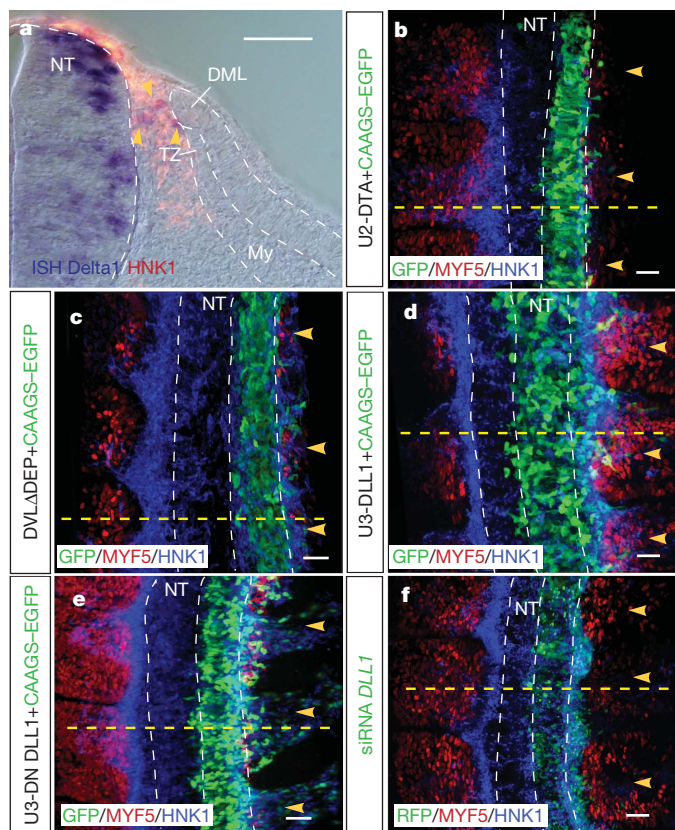


Fig. 8d; $n = 10/13$). We then electroporated *DLL1* under the control of the neural-crest-specific promoter in the neural tube and verified that this resulted in the overexpression of *DLL1* protein in neural crest cells (Supplementary Fig. 9a–c). We observed a significant increase of MYF5 expression (Fig. 4d and Supplementary Fig. 8e; $n = 9/13$). Loss of *DLL1* function was achieved by electroporating a dominant-negative form of *DLL1* (DN *DLL1*)²⁶ and with siRNAs against chick *DLL1*. DN *DLL1* protein was expressed in neural crest cells (Supplementary Fig. 9k–m) and the siRNA construct efficiently reduced the endogenous *DLL1* mRNA (Supplementary Fig. 9n–p) and protein (Supplementary Fig. 9q–s) levels. Both the DN *DLL1* ($n = 17/17$) and the *DLL1* siRNA ($n = 10/11$) resulted in a significant reduction in MYF5 staining (Fig. 4e, f and Supplementary Fig. 8f, g). Overexpression of *DLL1* in neural crest resulted in a robust activation of chick *HES1* mRNA expression (Supplementary Fig. 9d–f) and of the NOTCH reporter activity in somites (67%; Supplementary Fig. 9i, j), whereas electroporation of DTA or DN *DLL1* led to a near loss of NOTCH reporter activity (1.6% and 1.8%, respectively, Supplementary Fig. 9g, h, j; controls: 11%, Fig. 1c, d; Supplementary Fig. 2l), strongly supporting the hypothesis that NOTCH ligands presented

Figure 4 | Neural crest regulates myogenesis in somites through NOTCH signalling. **a**, Mosaic expression of chick *DLL1* (in blue, yellow arrowheads) within the HNK1-positive (in red) neural crest population. **b–f**, Confocal stacks of neural tube, neural crest and somites in dorsal view, 24 h after electroporation of one half of the neural tube with U2-DTA (**b**), CAGGS-DVLADP (**c**), U3-DLL1 (**d**), U2-DN DLL1 (**e**) (see Methods for details) and siRNA against chick *DLL1* (**f**). In green (**f**) native RFP; in green (**b–e**) GFP immunostaining; in red, MYF5 and in blue, HNK1. Dotted lines in **b–f** indicate the level of transverse sections shown in Supplementary Fig. 8c–g. Scale bars, 50 μ m.

by neural crest cells modulate NOTCH signalling in somites. When DTA, DLL1 or DN DLL1 was electroporated into neural crest, MYOD expression was affected similarly to MYF5 (Supplementary Fig. 10), indicating that the two major molecular players of the early myogenic program are similarly regulated by DLL1 from neural crest. The proliferation of progenitors within the DML was not significantly changed in these experiments (Supplementary Fig. 11a–l), further supporting the hypothesis (Fig. 3t, u) that NOTCH signalling regulates the progressive differentiation of the muscle progenitor pool within the DML. Because neural crest emigrates from the neural tube during a limited time period (about 24 h from when migration initiates), the neural-crest-mediated regulation of muscle growth is limited to the initial phases of myotome formation. However, this may have long-term consequences on muscle growth, as we observed significant changes in myotome growth 48 h after electroporation of DTA, DLL1 or DN DLL1 into the neural crest, that is, 24 h after crest migration has ceased (Supplementary Fig. 12a–s). As controls, we verified that the neural crest manipulations did not affect the expression of the known modulators of myotome formation in the dorsal neural tube, that is, *WNT1*, *WNT3A* and *BMP4* (Supplementary Fig. 13a–l). It is unclear whether the same regulatory mechanisms are used in mouse. Hypomorph *DLL1* mouse mutants displayed an enlarged primary myotome²⁰. However, as *Dll1* is expressed in both paraxial mesoderm and neural crest in early mouse embryo, the source of Notch signalling that engenders this phenotype remains to be defined. To examine this question, the inhibition of *Dll1* activity in specific cell types of the mouse embryo will be required.

Our model suggests an additional role of the NOTCH pathway during myogenesis whereby, within a population of DML cells all exposed to uniform gradients of myogenic activating factors, only those cells that transiently activate the NOTCH pathway undergo myogenesis. Transient NOTCH signalling is triggered by the NOTCH ligand *DLL1* carried and presented by migrating neural crest cells in a 'kiss and run' mode of signalling transduction (Supplementary Movie 2). This links the timing of myotome formation to that of neural crest migration, providing a mechanistic link for the concurrence of these two events (Supplementary Fig. 14a–g). The ability of migrating cells to influence cell fate in neighbouring tissues may reveal a general principle for generating pulses of signal activation that result in the differentiation of a defined subset of cells within a stem or progenitor pool.

METHODS SUMMARY

Electroporation, vectors, time-lapse experiments and confocal analyses. Further details can be found in Methods. The somite electroporation technique has been described elsewhere^{6,27}. Time-lapse experiments were performed essentially as described²⁷ on transverse slices of embryos.

Quantifications and statistical analyses. On average, more than 2,300 cells were counted per point to compute the corresponding quantifications shown in Figs 1–3 and Supplementary Figs 2–6. Statistical analyses were performed using the GraphPad Prism software.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 September 2010; accepted 24 February 2011.

Published online 15 May 2011.

- Shimojo, H., Ohtsuka, T. & Kageyama, R. Oscillations in Notch signaling regulate maintenance of neural progenitors. *Neuron* **58**, 52–64 (2008).
- Joubin, K. & Stern, C. D. Molecular interactions continuously define the organizer during the cell movements of gastrulation. *Cell* **98**, 559–571 (1999).
- Palmeirim, I., Henrique, D., Ish-Horowicz, D. & Pourquie, O. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91**, 639–648 (1997).

- Denetclaw, W. F. Jr, Berdugo, E., Venters, S. J. & Ordahl, C. P. Morphogenetic cell movements in the middle region of the dermomyotome dorsomedial lip associated with patterning and growth of the primary epaxial myotome. *Development* **128**, 1745–1755 (2001).
- Venters, S. J. & Ordahl, C. P. Persistent myogenic capacity of the dermomyotome dorsomedial lip and restriction of myogenic competence. *Development* **129**, 3873–3885 (2002).
- Gros, J., Scaal, M. & Marcelle, C. A two-step mechanism for myotome formation in chick. *Dev. Cell* **6**, 875–882 (2004).
- Kahane, N., Cinnamon, Y. & Kalcheim, C. The cellular mechanism by which the dermomyotome contributes to the second wave of myotome development. *Development* **125**, 4259–4271 (1998).
- Kahane, N., Cinnamon, Y. & Kalcheim, C. The roles of cell migration and myofiber intercalation in patterning formation of the postmitotic myotome. *Development* **129**, 2675–2687 (2002).
- Cinnamon, Y., Kahane, N. & Kalcheim, C. Characterization of the early development of specific hypaxial muscles from the ventrolateral myotome. *Development* **126**, 4305–4315 (1999).
- Kahane, N., Cinnamon, Y., Bachelet, I. & Kalcheim, C. The third wave of myotome colonization by mitotically competent progenitors: regulating the balance between differentiation and proliferation during muscle development. *Development* **128**, 2187–2198 (2001).
- Ben-Yair, R. & Kalcheim, C. Lineage analysis of the avian dermomyotome sheet reveals the existence of single cells with both dermal and muscle progenitor fates. *Development* **132**, 689–701 (2005).
- Gros, J., Manceau, M., Thome, V. & Marcelle, C. A common somitic origin for embryonic muscle progenitors and satellite cells. *Nature* **435**, 954–958 (2005).
- Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**, 948–953 (2005).
- Kassar-Duchossoy, L. et al. Pax3/Pax7 mark a novel population of primitive myogenic cells during development. *Genes Dev.* **19**, 1426–1431 (2005).
- Hirsinger, E. et al. Notch signalling acts in postmitotic avian myogenic cells to control MyoD activation. *Development* **128**, 107–116 (2001).
- Fryer, C. J., Lamar, E., Turbachova, I., Kintner, C. & Jones, K. A. Mastermind mediates chromatin-specific transcription and turnover of the Notch enhancer complex. *Genes Dev.* **16**, 1397–1411 (2002).
- Weng, A. P. et al. Growth suppression of pre-T acute lymphoblastic leukemia cells by inhibition of notch signaling. *Mol. Cell. Biol.* **23**, 655–664 (2003).
- Das, R. M. et al. A robust system for RNA interference in the chicken using a modified microRNA operon. *Dev. Biol.* **294**, 554–563 (2006).
- Vasyutina, E., Lenhard, D. C. & Birchmeier, C. Notch function in myogenesis. *Cell Cycle* **6**, 1450–1453 (2007).
- Schuster-Gossler, K., Cordes, R. & Gossler, A. Premature myogenic differentiation and depletion of progenitor cells cause severe muscle hypotrophy in *Delta1* mutants. *Proc. Natl Acad. Sci. USA* **104**, 537–542 (2007).
- Vasyutina, E. et al. *RBP-J (Rbpsi)* is essential to maintain muscle progenitor cells and to generate satellite cells. *Proc. Natl Acad. Sci. USA* **104**, 4443–4448 (2007).
- De Calisto, J., Araya, C., Marchant, L., Riaz, C. F. & Mayor, R. Essential role of non-canonical Wnt signalling in neural crest migration. *Development* **132**, 2587–2597 (2005).
- Wallingford, J. B. et al. Dishevelled controls cell polarity during *Xenopus* gastrulation. *Nature* **405**, 81–85 (2000).
- Rothbacher, U. et al. Dishevelled phosphorylation, subcellular localization and multimerization regulate its role in early embryogenesis. *EMBO J.* **19**, 1010–1022 (2000).
- Gros, J., Serrallbo, O. & Marcelle, C. WNT11 acts as a directional cue to organize the elongation of early muscle fibres. *Nature* **457**, 589–593 (2009).
- Henrique, D. et al. Expression of a *Delta* homologue in prospective neurons in the chick. *Nature* **375**, 787–790 (1995).
- Rios, A. C., Denans, N. & Marcelle, C. Real-time observation of Wnt β -catenin signaling in the chick embryo. *Dev. Dyn.* **239**, 346–353 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank N. Rosenthal and P. Currie for critical reading of the manuscript. This study was funded by grants from the Agence Nationale pour la Recherche (ANR), and by the EU 6th Framework Programme Network of Excellence MYORES. The help of P. Weber, S. Firth, C. Johnson and I. Harper from Imaging Facilities (IBDML, Marseille and MMI, Monash University) is acknowledged.

Author Contributions A.C.R. and C.M. conceived the experiments. A.C.R. predominantly performed the work with the help of O.S. D.S. designed the animation. C.M. supervised the project and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.M. (christophe.marcelle@monash.edu).

METHODS

Electroporation and confocal analysis. The somite electroporation technique that was used throughout this study has been described elsewhere^{6,27}. Briefly, we targeted the expression of various constructs to the dorsomedial portion of newly formed interlimb somites of Hamburger–Hamilton (HH) stage 15–16 chick embryos (24–28 somite)²⁸. We have previously shown that this technique allows the specific expression of cDNA constructs in the DML of the dermomyotome⁶. To target the neural crest population, we electroporated the dorsal neural tube of HH stage 13–14 chick embryos at the level of the presomitic mesoderm.

The following constructs have been previously published: HES1–d2EGFP and the HES1–EGFP²⁹ (provided by R. Kageyama) contain the mouse *Hes1* promoter region upstream of destabilized or stable GFP. The CAGGS–H2B–RFP (provided by S. Tajbakhsh) contains a fusion of histone 2B with RFP downstream of the CAGGS strong ubiquitous promoter (CMV/chick β -actin promoter/enhancer). The CAGGS–EGFP³⁰ contains the CAGGS promoter followed by the EGFP reporter gene. The pCAB–HA–NICD–IRES–GFP (provided by N. Daudet) contains an HA-tagged NICD under the control of the CAGGS promoter³¹. The doxycyclin inducible system is composed of two plasmids that are co-electroporated: first, the pCIRX–rtTA–IRES–DsRed³² (provided by O. Pourqu  ) contains a Tet-On Advanced transactivator (rtTA, Clontech) downstream of the CAGGS promoter. The IRES–DsRed–Express allows the detection of electroporated cells. Second, the pBI–HANICD–EGFP is the response plasmid (Clontech) in which the HA-tagged constitutively active form of NOTCH, NICD, was cloned. The bidirectional tetracycline-response element drives the expressions of EGFP (which serves as an internal control of the induced response, see Supplementary Fig. 2a, b) and HANICD. pCLGFP–DVLADep contains a mutated form of *Xenopus* Dishevelled that lacks the DEP domain, driven by the CAGGS promoter²⁵. This construct contains also EGFP driven by its own SV40 promoter. The siRNA directed against chick *NOTCH1* has been described elsewhere¹⁸.

We made new constructs for this study: to construct the HES1 nVENUS-PEST, a destabilized nuclear Venus GFP variant³³ was inserted downstream of the mouse *Hes1* promoter region²⁹. The CAGGS–DN MAML1–EGFP contains a truncated, dominant-negative form of the human Mastermind (DN MAML1), fused with EGFP¹⁷ downstream of the CAGGS promoter. The pCAB–HA–NICD was constructed by removing the EGFP reporter from pCAB–HA–NICD–IRES–GFP. The U2- and U3-EGFP were made by inserting the U2 and U3 evolutionary conserved *Sox10* enhancer sequences³⁴ in the TK–EGFP³⁵ plasmid, that contains the thymidine kinase minimal promoter upstream of the EGFP. The diphtheria toxin gene³⁶, the chick *DLL1* or a dominant-negative form of this gene³⁷ were inserted in the U2 or the U3–TK–EGFP in place of the EGFP to obtain the U2–DTA, the U2–DNDLL1 and the U3–DLL1 electroporation vectors. To detect electroporated cells, those plasmids were electroporated with a pCAGGS–EGFP. We have constructed two RNA interference plasmids as described previously¹⁸ that each express two siRNAs directed against chick *DLL1*. Sequences TCACAGCGATA ACTCCGATAAAA and TGCAGGAGTTTGTCACCAAGAA were inserted in siRNA chick *DLL1* A, whereas sequences GATTTCAGTATATCCACTTCAA and CCGGCACCTTCTCGCTCATCAT were inserted in siRNA chick *DLL1* B. Electroporation of plasmids A, B, or A together with B efficiently decreased the endogenous expression of chick *DLL1* mRNA and protein, whereas the electroporation of siRNA directed against luciferase had no effect on chick *DLL1* expression. An RFP reporter gene is inserted in the same constructs to detect electroporated cells.

Antibody stainings and BrdU labelling. For BrdU labelling, embryos were incubated for 30 min with 50 μ l of a 1 mg ml^{−1} BrdU (Sigma) solution. Whole-mount antibody stainings were performed as described²⁵. The following antibodies were used: rabbit polyclonals directed against chick myogenic regulatory factors MYF5 and MYOD³⁸; chick *DLL1*³⁹; and anti-RFP (Abcam); chicken polyclonals against EGFP (Abcam); rat polyclonals against the HA tag and anti-BrdU (Abcam). We also used monoclonals against the dermomyotome and dorsal neural tube marker PAX7 and against terminal myogenic differentiation marker MyHC (MF20) (obtained from the Developmental Studies Hybridoma Bank); and the neural-crest-specific monoclonal antibody HNK1 (provided by A. Eichmann).

In situ hybridization. The following probes were used: chick HES1/cHairy2 (ref. 40) and chick²⁷ *DLL1* and chick LFNG (provided by O. Pourqu  ), and 400 bp cDNA clones coding for fragments of chick *WNT1*, *WNT3A* and a 1 kb chick *BMP4* probe⁴¹.

Doxycyclin-mediated induction of NOTCH signalling. Eight hours after electroporation of pCIRX–rtTA–IRES–DsRed and pBI–HANICD–EGFP, doxycyclin (300 μ l of a 0.1 μ g ml^{−1} solution) was added onto the embryos, and it was either washed off after one hour with PBS for transitory upregulation of NICD, or left overnight, for permanent expression of this molecule. We verified that the response plasmid is completely silent before doxycyclin addition (that is, no EGFP expression, Supplementary Fig. 6a) while it is strongly and rapidly activated 6 h after doxycyclin addition (Supplementary Fig. 6b).

Time-lapse experiments and confocal analyses. Time-lapse experiments were performed essentially as described²⁷ on transverse slices (250 μ m) of embryos. Embryo slices were filmed for 11 h at 37 °C with a confocal inverted Leica SP5 microscope equipped with a resonant scanner, at the rate of one image stack per ten minutes. Confocal images were acquired transversally over a thickness of 100 μ m; Supplementary Movie 1 corresponds to a fraction (10 μ m thick) of the acquired images. Dorsal views shown in Figs 1–4 are projections of stacks of confocal images. Confocal stacks of images were visualized and analysed with the Imaris software suite. Cell countings were performed using the Improvision Velocity software suite.

Quantifications and statistical analyses. Electroporation results in the transfection of a portion of the targeted cell population, which is variable from embryo to embryo. To precisely evaluate the phenotypes obtained after electroporation of cell-autonomously acting cDNA constructs, the number of positive cells was compared to the total number of electroporated cells, recognized by an internal fluorescent reporter construct. On average, more than 2,300 cells were counted per point and the corresponding quantifications are shown in Figs 1–3 and Supplementary Figs 2–6. This mode of quantification could not be applied when constructs were electroporated in one tissue while the effects were evaluated in another, such as in experiments shown in Fig. 4 and Supplementary Figs 8–11. In this case, we report the number of embryos in which we observed a phenotype similar to the one that is illustrated in the figures, over the total number of electroporated embryos. Statistical analyses were performed using the GraphPad Prism software. Mann–Whitney non-parametric two-tail testing was applied to populations to determine the *P* values indicated in the figures. In each graph, columns correspond to the mean and error bars correspond to the standard deviation. ****P* value < 0.0001.

28. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *Dev. Dyn.* **195**, 231–272 (1992).
29. Ohtsuka, T. et al. Visualization of embryonic neural stem cells using *Hes* promoters in transgenic mice. *Mol. Cell. Neurosci.* **31**, 109–122 (2006).
30. Tobiume, M. et al. Inefficient enhancement of viral infectivity and CD4 downregulation by human immunodeficiency virus type 1 Nef from Japanese long-term nonprogressors. *J. Virol.* **76**, 5959–5965 (2002).
31. Daudet, N. & Lewis, J. Two contrasting roles for Notch activity in chick inner ear development: specification of prosensory patches and lateral inhibition of hair-cell differentiation. *Development* **132**, 541–551 (2005).
32. Imura, T. & Pourqu  , O. Collinear activation of *Hoxb* genes during gastrulation is linked to mesoderm cell ingression. *Nature* **442**, 568–571 (2006).
33. Nagoshi, E. et al. Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell* **119**, 693–705 (2004).
34. Werner, T., Hammer, A., Wahlbuhl, M., Bosl, M. R. & Wegner, M. Multiple conserved regulatory elements with overlapping functions determine *Sox10* expression in mouse embryogenesis. *Nucleic Acids Res.* **35**, 6526–6538 (2007).
35. Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y. & Kondoh, H. Functional analysis of chicken *Sox2* enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev. Cell* **4**, 509–519 (2003).
36. Maxwell, I. H., Maxwell, F. & Glode, L. M. Regulated expression of a diphtheria toxin A-chain gene transfected into human cells: possible strategy for inducing cancer cell suicide. *Cancer Res.* **46**, 4660–4664 (1986).
37. Henrique, D. et al. Maintenance of neuroepithelial progenitor cells by Delta–Notch signalling in the embryonic chick retina. *Curr. Biol.* **7**, 661–670 (1997).
38. Manceau, M. et al. Myostatin promotes the terminal differentiation of embryonic muscle progenitors. *Genes Dev.* **22**, 668–681 (2008).
39. Henrique, D. et al. *cash4*, a novel achaete-scute homolog induced by Hensen's node during generation of the posterior nervous system. *Genes Dev.* **11**, 603–615 (1997).
40. Jouve, C. et al. Notch signalling is required for cyclic expression of the hairy-like gene *HES1* in the presomitic mesoderm. *Development* **127**, 1421–1429 (2000).
41. Marcelle, C., Stark, M. R. & Bronner-Fraser, M. Coordinate actions of BMPs, Wnts, Shh and noggin mediate patterning of the dorsal somite. *Development* **124**, 3955–3963 (1997).

Non-adaptive origins of interactome complexity

Ariel Fernández^{1,2} & Michael Lynch³

The boundaries between prokaryotes, unicellular eukaryotes and multicellular eukaryotes are accompanied by orders-of-magnitude reductions in effective population size, with concurrent amplifications of the effects of random genetic drift and mutation¹. The resultant decline in the efficiency of selection seems to be sufficient to influence a wide range of attributes at the genomic level in a non-adaptive manner². A key remaining question concerns the extent to which variation in the power of random genetic drift is capable of influencing phylogenetic diversity at the subcellular and cellular levels^{2–4}. Should this be the case, population size would have to be considered as a potential determinant of the mechanistic pathways underlying long-term phenotypic evolution. Here we demonstrate a phylogenetically broad inverse relation between the power of drift and the structural integrity of protein subunits. This leads to the hypothesis that the accumulation of mildly deleterious mutations in populations of small size induces secondary selection for protein–protein interactions that stabilize key gene functions. By this means, the complex protein architectures and interactions essential to the genesis of phenotypic diversity may initially emerge by non-adaptive mechanisms.

Here we examine whether established gene orthologies reveal a role for drift in phylogenetic patterns of protein structural evolution. Although evolutionary change at the structural level is unlikely to destabilize greatly the native fold of an essential protein, as the complete loss of function would generally be unbearable, the drift hypothesis predicts a negative relation between population size (N) and the accumulation of mildly deleterious amino-acid substitutions. The following examination of the structures of orthologous proteins from vastly different lineages suggests that the enhanced power of drift in eukaryotes (multicellular species in particular) results in a qualitative reduction in the stability of protein–water interfaces (PWIs) through the partial exposure of paired backbone polar groups (amides and carbonyls) that are otherwise protected in prokaryotes. In effect, the reduced efficiency of selection in small- N species encourages the accumulation of mild structural deficiencies in the form of solvent-accessible backbone hydrogen bonds (SABHBs), which lead to protein structures that are more ‘open’ and vulnerable to fold-disruptive hydration (Fig. 1a) and create protein–water interfacial tension (PWIT; Supplementary Fig. 1)⁵ by hindering the hydrogen-bonding capabilities of nearby water molecules.

We argue that the emergence of unfavourable PWIs promotes the secondary recruitment of novel protein–protein associations that restore structural stability by reducing PWI. Under this hypothesis, complex organisms may frequently develop protein–protein interactions not as immediate vehicles for novel adaptive functions, but as compensatory mechanisms for retaining key gene functions. Once in place, such physical contact between interacting proteins may provide a selective environment for the further emergence of entirely novel protein–protein interactions underlying cellular and organismal complexities. Our suggestion that the hallmark of eukaryotic evolution, the origin of interactome complexity, may have arisen in part as a passive consequence of the enhanced power of drift reduces the need to invoke direct long-term selective advantages of phenotypic complexity⁶.

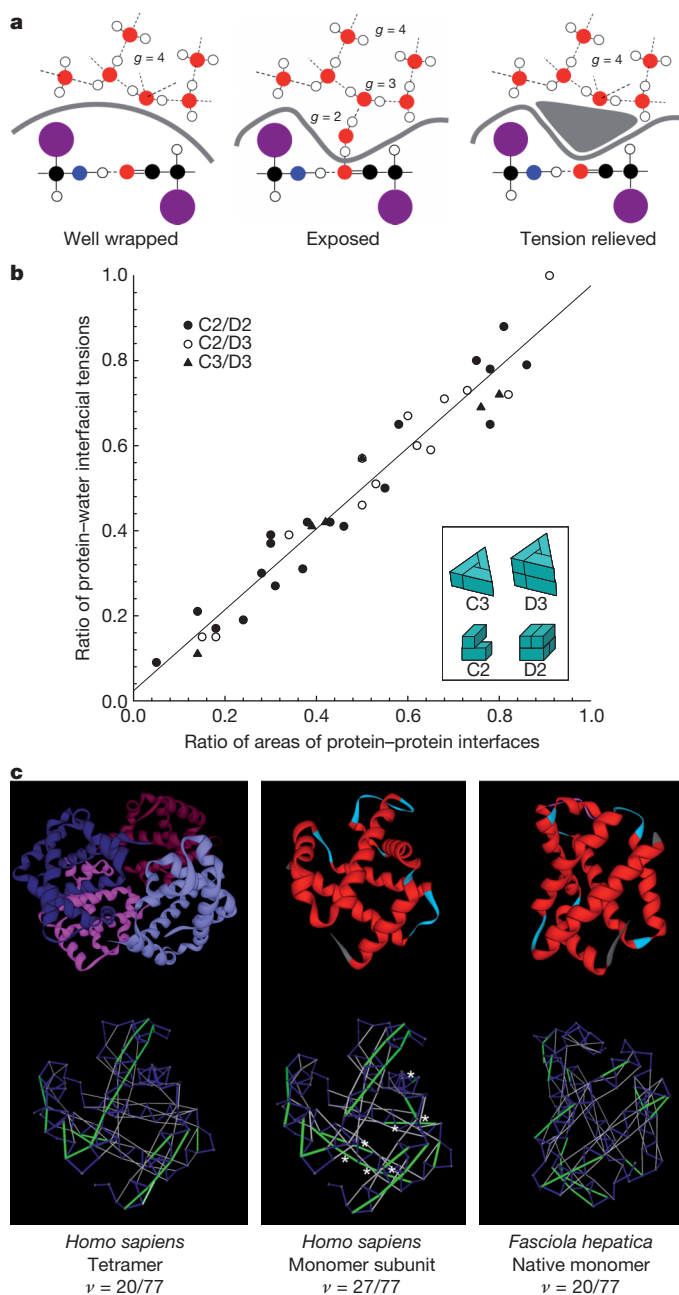
To gain insight into the evolution of interactome complexity, we derived quantitative measures of the PWIT as indicators of potential molecular interactivity. To estimate the PWIT of a protein, we computationally equilibrated the protein structure in surrounding water, using the function $g(\mathbf{r})$ to represent the time-averaged coordination (number of hydrogen bonds) associated with a water molecule at position \mathbf{r} (Fig. 1a), and integrating over the entire protein surface all water molecules within a 10 Å radius (the thickness of four layers of water molecules). Compared with bulk water (where $g = 4$), interfacial water molecules may have reduced hydrogen-bonding opportunities ($g < 4$) and often counterbalance these losses by interacting with polar groups on the protein surface. Thus the PWIT parameter integrates information on unfavourable local decreases in g and favourable polarization contributions from the protein to yield the free-energy cost, ΔG_{if} , of spanning the protein–water interface (Methods). A high PWIT signals a high propensity for protein–protein associations, which reduce the PWI area.

To validate the use of PWIT as a measure of interactivity, we examined an exhaustive catalogue of contact topologies for protein complexes with one to six subunits, with each topology being evaluated with one or more non-homologous complexes using structures in the Protein Data Bank (PDB) (Supplementary Table 1). For each complex, we computed the total protein–protein interface area after identifying the residues engaged in intermolecular contacts⁷. For each protein subunit, the protein–protein interface is contained within the PWI region that generates tension in the free subunit, and there is a tight correlation between the surface areas for both regions, implying that regions on the protein surface generating PWIT (i.e. those with $g < 4$ for nearby water) actually promote associations (Supplementary Figs 1 and 2a). Next, we verified that protein surface regions generating PWIT coincide with the affinity-contributing regions at protein–protein interfaces. To this end, we tested the value of PWIT as a promoter of protein associations by focusing on the interface for the 1:1 human growth hormone (hGH)–receptor complex⁸ (Supplementary Fig. 2b) for which the consequences of amino-acid substitutions have been extensively evaluated. Our analysis reveals a strong correlation between the change in PWIT induced by site-specific mutagenesis of interfacial residues and the association free-energy difference created by the alteration of the hormone–receptor interface (Supplementary Fig. 2c).

Comparison of orthologous proteins engaging in different levels of homo-oligomerization in different species⁹ further supports the view that PWIT serves as a measure of the propensity for protein–protein association. The ratio of protein–protein interface areas (lower to higher degrees of complexation; Supplementary Table 2) exhibits a strong positive correlation with the ratio of PWITs for the respective free subunits (Fig. 1b). As complexes with higher degrees of oligomerization arise from lower-order complexes, this implies that the degree of cooperativity among subunits correlates with the PWIT of the basic subunit.

Hydrophobic regions on protein surfaces obviously contribute to PWIT, but analysis of proteins exhibiting association propensity (Supplementary Table 2) shows that the regions generating $73 \pm 5\%$

¹Department of Computer Science, The University of Chicago, Chicago, Illinois 60637, USA. ²Department of Bioengineering, Rice University, Houston, Texas 77005, USA. ³Department of Biology, Indiana University, Bloomington, Indiana 47405, USA.



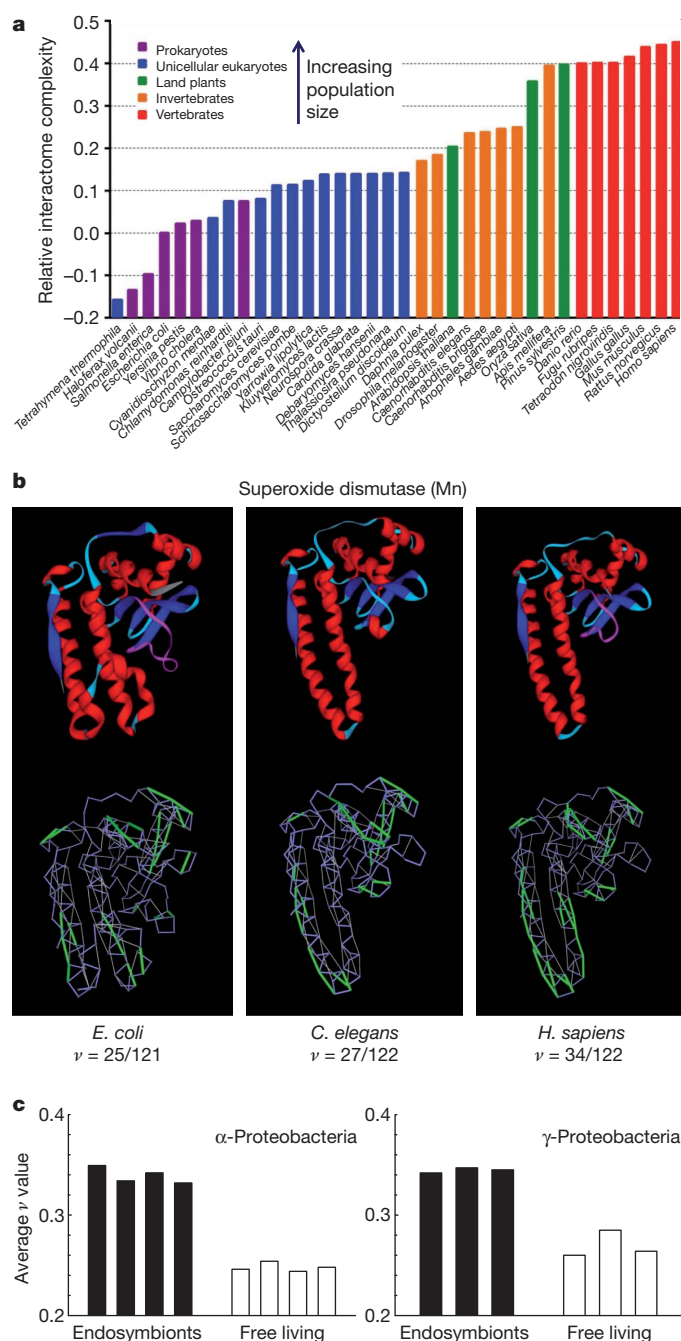
of the PWIT (Supplementary Information and Supplementary Fig. 1) arise from SABHBs. The resultant hydration of backbone polar groups (amides and carbonyls) causes a loss of coordination for local water molecules, which increases surface tension and creates an unstable PWI, as the cavities cannot accommodate a bulk-like water molecule¹⁰. As an example of how such a structural deficiency can be alleviated through a protein association, an isolated α -subunit of the human haemoglobin tetramer has seven SABHBs that become protected within the tetrameric complex, such that the ratio (v) of SABHBs to total BHBs in the complex-associated subunit is the same as that in the natively monomeric unit for haemoglobin from the trematode *Fasciola hepatica* (Fig. 1c).

To evaluate whether the accumulation of structural deficiencies of proteins is generally encouraged by random genetic drift, and in turn enhances the propensity for establishing protein complexes, we examined a set of 106 orthologous water-soluble proteins (sequence identity greater than 30%)^{11,12} with PDB-reported structures for at least two species. We considered 36 species with vastly different population sizes¹², each containing proteins in at least 90 of the 106 orthologous groups

Figure 1 | Structural deficiencies in soluble proteins promote protein associations. **a**, Hydration of exposed polar backbone induces interfacial tension by causing water molecules near the defect to relinquish part of their coordination ($g < 4$) relative to the level in surrounding bulk solvent ($g = 4$). White represents hydrogen atoms; red, oxygen; blue, nitrogen; black, carbon; the larger purple circles denote side chains for amino acids. Hydrogen bonds are denoted by dashed lines. Thick grey lines outline the external surface of the overall protein molecule, and the underlying structure represents two amino acids made adjacent by the protein architecture and bound by a hydrogen bond between the backbone amide (blue:white) of one amino acid and carbonyl (red:black) of the other. Water molecules are shown as angular red and white segments, with the coordination number g denoting the number of hydrogen bonds associated with a water molecule ($g = 4$ for bulk water; $g < 4$ for confined interfacial water). In the centre, the structure of the protein causes local exposure and unfavourable hydration of the polar backbone, whereas the absence of such local interactions between water molecules and the well-wrapped proteins on the left and right reduces interfacial tension (interfacial water is bulk-like, retaining the maximum coordination $g = 4$). **b**, Comparison of orthologous proteins with different levels of homo-oligomerization reveals that the PWIT is an indicator of the propensity for cooperative improvement/refinement of protein function through complexation. The ratio of protein-protein interfaces (small to large) was determined for pairs of orthologous proteins with different levels of oligomerization in different species (Supplementary Table 2) and plotted against the ratio of PWITs for the respective free subunits. The tight correlation ($r^2 = 0.94$) reveals that interspecific differences in PWIT accompany differences in levels of oligomerization, thus providing a measure of potential allosteric or cooperative improvement of basic protein function. Complexes with cyclic rotational symmetry (C2, C3, ...) can further oligomerize into complexes with dihedral (D2, D3, ...) symmetry, as shown in the idealized diagrams in the lower right. For example, C2 complexes can dimerize into D2 complexes, trimerize into D3 complexes, etc., whereas a D3 complex can also be obtained by dimerization of a C3 complex. For the protein-protein interface and PWIT ratios examined, the interface for the subunit in the complex with lower-order symmetry is compared with that in the complex with higher-order symmetry, yielding analyses based on protein pairs contrasted within three groupings: C2 versus D2, C2 versus D3, and C3 versus D3. **c**, The SABHB patterns from two haemoglobins with different oligomerization levels in their native states are compared. In the bottom panels, the protein backbone is represented by virtual bonds in blue joining α -carbons, with well-protected BHBs shown as light grey and SABHBs as green lines joining the α -carbons of the paired residues. The ribbon representations of the human complex and dissociated subunit (chain A in PDB.2DN2, left and centre, respectively) are included as aids to the eye, representing the structuring of the backbone in each subunit. The free subunit isolated from the tetramer in *H. sapiens* (PDB.2DN2, chain A, centre) has seven excess SABHBs (denoted by stars) when compared with the subunit within the tetrameric complex, where they are well-protected intermolecularly, alleviating interfacial tension. As a consequence of this better wrapping, the overall extent of structural deficiency (v value) for the subunit within the human complex is identical to that of the natively monomeric haemoglobin from the trematode *F. hepatica* (PDB.2VYW). This raises the possibility that the accumulation of structural deficiencies in the mammalian haemoglobin subunit promoted the emergence of an oligomeric association as a means of reducing excess interfacial tension. The structural displays were obtained by uploading the PDB text files into the program YAPview, a display of local backbone desolvation of soluble proteins that can be downloaded from the link 'Dehydron Calculator' at <http://www.owl.net.rice.edu/~arifer/>.

(Supplementary Tables 3–5). Template-based three-dimensional structures for orthologues lacking PDB-reported structures were constructed by homology threading^{13,14}, and evaluated, ranked and selected according to the energetic proximity between template and model¹⁵. The accuracy of this homology-based prediction of PWIT was determined with a test set of proteins with PDB-reported structures from two species, subjecting one member of each orthologous pair to homology threading through the other. Comparison of the indirect and direct estimates of PWIT demonstrates that when sequence identities are greater than 35%, the predicted PWIT diverges less than 10% from the more direct estimate for the same protein (Supplementary Fig. 3).

For each protein structure, $g(r)$ was obtained as described in Methods, and the relative propensities for protein association across orthologues were then determined by assessing differences in the free-energy cost



ΔG_{if} among species. We estimated the relative complexation propensity $M_{j,n}$ of a protein in orthologue group j ($1, \dots, 106$) from species n ($1, \dots, 36$) by adopting *Escherichia coli* as a reference species ($n = 1$):

$$M_{j,n} = [(\Delta G_{if})_{j,n} - (\Delta G_{if})_{j,1}] / (\Delta G_{if})_{j,1} \quad (1)$$

With this index, $M_{j,1} = 0$ for all proteins in *E. coli*, and taxa with less well-wrapped proteins (and hence greater propensity for complexation), have positive values.

The mean value of species-specific estimates of $M_{j,n}$ over all proteins evaluated is negatively correlated with the approximate effective population sizes of species (Fig. 2a), given that the average ranking of the latter is prokaryotes > unicellular eukaryotes > invertebrates > vertebrates and land plants^{1,2}. A specific example of a trend towards increasing structural openness with reduced population sizes is illustrated in Fig. 2b, where the SABHB patterns and v values for orthologues of the enzyme superoxide dismutase are compared across three species.

Figure 2 | Structural degradation enhances PWIT and promotes protein interactivity in species with small population sizes. **a**, Potential for interactome complexity of 36 species with diverse population sizes (Supplementary Table 2), relative to *E. coli*. To highlight the relative power of random genetic drift, bars are colour-coded to reflect groupings of species in broad population-size categories. **b**, Overall structural deficiency of orthologues of the enzyme superoxide dismutase (Mn), revealing a progressive accumulation of SABHBs in the orthologues of the bacterium *E. coli*, the nematode *Caenorhabditis elegans* and *H. sapiens*. The upper ribbon representations illustrate the structural conservation across orthologues (respective PDB accession numbers 3ot7, 3dc6, 2adq). The conventional colour coding is red, blue, magenta and light blue for helix, β -strand, loop and turn, respectively. **c**, Average structural deficiency (v value) of protein orthologues for intracellular and free-living bacterial species. Species identities, progressing from left to right are as follows: α -Proteobacteria—*Rickettsia typhi*, *Orientia tsutsugamushi*, *Anaplasma centrale* str. Israel, *Wolbachia* sp. wRi, *Rhodospirillum centenum* SW, *Magnetospirillum magneticum*, *Silicibacter TM1040*, *Erythrobacter litoralis*; γ -Proteobacteria—*Buchnera aphidicola*, *Wigglesworthia brevipalpis*, *Candidatus Blochmannia pennsylvanicus*, *Marinomonas MWYL1*, *E. coli*, *Pseudomonas aeruginosa*. Only proteins with orthologues across the full set of species within each group were considered for analysis (Supplementary Tables 6 and 7).

The results from Fig. 2a and an additional analysis (Supplementary Fig. 4) support the hypothesis that large organisms with small population sizes experience a significant enough increase in the power of random genetic drift to magnify the accumulation of mild structural deficiencies in the form of SABHBs, resulting on average in proteins with a more solvent-exposed or 'open' structure. By contrast, mutations to SABHBs are more frequently excluded by selection in species with larger population sizes (for example, prokaryotes). Thus, because SABHBs are the main determinants of interfacial tension (Supplementary Fig. 1), the proteins of large organisms have a greater inherent tendency to form novel protein–protein associations (Fig. 1a). This suggests that increases in protein-network complexity in multicellular species may in part owe their origins to modifications to the intracellular selective environment induced by non-adaptive structural degradation of individual proteins.

One concern with the preceding interpretation is the order of events: does an initial degradation of architectural integrity of individual proteins in response to random genetic drift induce secondary selection for the recruitment of interacting partners, or does the emergence of cellular complexity (and increased protein interactivity) precede secondary changes in protein sequence to accommodate such interactions? One way to evaluate this matter is to compare proteins from related species that have experienced relatively recent divergences in effective population sizes but no major modifications in intracellular complexity or emergence of multicellularity.

To achieve this task, we compared orthologous genes from endosymbiotic/intracellular bacteria and their free-living relatives, as the former are thought to have experienced substantial reductions in effective population sizes¹⁶. Previous suggestions that intracellular bacteria experience elevated levels of random genetic drift have been based on ratios of substitution rates at silent and replacement sites, which can be biased indicators of the efficiency of selection if there is selection on silent sites. Although the lack of protein structural information for endosymbiotic species requires a sequence-based identification of SABHBs derived from reliable scores of native disorder propensity (Methods), the resultant analyses are broadly consistent with the hypothesis that an increase in the power of drift in microbes encourages the accumulation of structural defects in protein architecture (Fig. 2c). Free-living species, with larger effective population sizes, have consistently smaller v values for orthologous genes in both α - and γ -Proteobacteria. (Application of the same sort of analysis of disorder propensity across a set of 105 species and 541 proteins corroborates this result (Supplementary Figs 5–7).)

Taken together, our analyses support the hypothesis that the range of population sizes experienced by natural populations is sufficient to

induce significantly different patterns of evolution at the level of protein architecture. The resultant changes in the intracellular environment in small-*N* species provides an opportunity for the recruitment of stabilizing protein–protein interactions, yielding a plausible mechanism for the emergence of molecular complexities before their exploitation in phenotypic divergence^{9,17}. This hypothesis does not deny a potentially significant role for natural selection in using such novelties subsequent to their establishment, nor does it deny the fact that intramolecular compensatory mutations can alleviate some structural defects associated with SABHBs. However, our results do raise questions about the necessity of invoking an intrinsic advantage to organismal complexity, and provide a strong rationale for expanding comparative studies in molecular evolution beyond linear sequence analysis to evaluations of molecular structure.

METHODS SUMMARY

We determined the propensity of proteins to be engaged in associations that reduce the PWI by computing the PWIT. This thermodynamic parameter gives ΔG_{if} , the free-energy cost of spanning the PWI. The PWIT is computed as

$$\Delta G_{if} = \frac{1}{2} \{ a |\nabla g|^2 - |\mathbf{P}[g(\mathbf{r})]|^2 \} d\mathbf{r}, \quad (2)$$

where the term $\frac{1}{2} a |\nabla g|^2$, with $a = 9.02 \text{ mJ m}^{-1}$ at $T = 298 \text{ K}$ (Methods), accounts for tension-generating reductions in water coordination, and the polarization $\mathbf{P}[g(\mathbf{r})]$ accounts for dipole–electrostatic field interactions (Methods). For a given protein structure or template-based structural model, the field $g = g(\mathbf{r})$ used in the numerical integration of equation (2) was determined by equilibrating the water-embedded structure within an isothermal–isobaric (NPT) ensemble (with fixed parameters N = number of particles, P = pressure and T = temperature; Methods)^{10,18,19}. From structural coordinates, we determined the structural deficiencies (SABHBs)²⁰ that generate $73 \pm 5\%$ of the PWIT (Supplementary Information). We examined 106 groups of orthologous proteins identified using OrthoMCL^{11,12} for which there are PDB representatives from at least two species (usually *E. coli* and *Homo sapiens*, Supplementary Tables 3–5). We considered 36 representative species, each containing proteins in at least 90 of the 106 orthologue groups. Template-based three-dimensional structures for orthologues lacking a PDB-reported structure¹⁴ were constructed using MODELLER¹³, with side chains directly positioned with SCWRL²¹. The template and resulting model were evaluated, ranked and finally selected using ProSA¹⁵. The accuracy of homology models is shown in Supplementary Fig. 3. In cases where orthologous structural templates were unavailable, like the comparison of endosymbionts with free-living species, a sequence-based inference of SABHBs was performed based on an established anti-correlation between backbone protection and disorder propensity (Supplementary Fig. 5)²². The cross validation of homology- and disorder-based estimations of v values is given in Supplementary Fig. 6.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 August 2010; accepted 9 March 2011.

Published online 18 May 2011.

1. Lynch, M. & Conery, J. S. The origins of genome complexity. *Science* **302**, 1401–1404 (2003).
2. Lynch, M. *The Origins of Genome Architecture* (Sinauer, 2007).

3. Stoltzfus, A. On the possibility of constructive neutral evolution. *J. Mol. Evol.* **49**, 169–181 (1999).
4. Gray, M. W., Lukes, J., Archibald, J. M., Keeling, P. J. & Doolittle, W. F. Cell biology. Irremediable complexity? *Science* **330**, 920–921 (2001).
5. Rowlinson, J. S. & Widom, B. *Molecular Theory of Capillarity* (Oxford Univ. Press, 1982).
6. Lynch, M. The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc. Natl Acad. Sci. USA* **104** (Suppl.), 8597–8604 (2007).
7. Levy, E. D., Pereira-Leal, J. B., Chothia, C. & Teichmann, S. A. 3D Complex: a structural classification of protein complexes. *PLOS Comput. Biol.* **2**, e155 (2006).
8. Clackson, T., Ultsch, M. H., Wells, J. A. & de Vos, A. M. Structural and functional analysis of the 1:1 growth hormone:receptor complex reveals the molecular basis for receptor affinity. *J. Mol. Biol.* **277**, 1111–1128 (1998).
9. Levy, E. D., Boeri Erba, E., Robinson, C. V. & Teichmann, S. A. Assembly reflects evolution of protein complexes. *Nature* **453**, 1262–1265 (2008).
10. Fenimore, P. W., Frauenfelder, H., McCammon, B. H. & Young, R. D. Bulk solvent and hydration-shell fluctuations, similar to α - and β -fluctuations in glasses, control protein motions and functions. *Proc. Natl Acad. Sci. USA* **101**, 14408–14413 (2004).
11. Ostlund, G. et al. InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.* **38**, D196–D203 (2010).
12. Gabaldon, T. et al. Joining forces in the quest for orthologs. *Genome Biol.* **10**, 403 (2009).
13. Sali, A. & Blundell, T. L. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
14. Zhou, H. & Skolnick, J. Improving threading algorithms for remote homology modeling by combining fragment and template comparisons. *Proteins* **78**, 2041–2048 (2010).
15. Wiederstein, M. & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407–W410 (2007).
16. Moran, N. A. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc. Natl Acad. Sci. USA* **93**, 2873–2878 (1996).
17. Kuriyan, J. & Eisenberg, D. The origin of protein interactions and allostery in colocalization. *Nature* **450**, 983–990 (2007).
18. Rizzo, R. C. & Jorgensen, W. L. OPLS All-atom model for amines: resolution of the amine hydration problem. *J. Am. Chem. Soc.* **121**, 4827–4836 (1999).
19. Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
20. Fernández, A. & Berry, R. S. Golden rule for buttressing vulnerable soluble proteins. *J. Proteome Res.* **9**, 2643–2648 (2010).
21. Canutescu, A. A., Shelenkov, A. & Dunbrack, R. L. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci.* **12**, 2001–2014 (2003).
22. Pietrosoli, N., Crespo, A. & Fernández, A. Dehydration propensity of order-disorder intermediate regions in soluble proteins. *J. Proteome Res.* **6**, 3519–3526 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements A.F. was supported by National Institutes of Health grant R01GM072614, and by the Institute of Biophysical Dynamics and the Department of Computer Science at The University of Chicago. M.L. was supported by National Institutes of Health grant R01GM036827 and National Science Foundation grant EF-0827411.

Author Contributions A.F. and M.L. conceived the project and wrote the paper. A.F. collected the orthologue groups across 36 species with sufficient structural representation, performed the structural analysis and determined the interaction propensities across orthologues.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.L. (milynch@indiana.edu) or A.F. (ariel@uchicago.edu).

METHODS

Computation of PWIT. The parameter a in equation (2) is obtained from the interfacial tension of a large non-polar sphere with radius θ in the limit $\theta/1 \text{ nm} \rightarrow \infty$. Thus we get $a = 9.02 \text{ mJ m}^{-1} = \lim_{\theta/1 \text{ nm} \rightarrow \infty} [\gamma/(4\pi\theta^2)]^{1/2} |\nabla g|^2 d\mathbf{r}$, where $\gamma = 72 \text{ mJ m}^{-2}$ is the bulk surface tension of water at 298 K, and $|\nabla g|^2 d\mathbf{r} = O(\theta^2)$ since $\nabla g \neq 0$ only in the vicinity of the interface. To determine the g -dependence of polarization $\mathbf{P} = \mathbf{P}(\mathbf{r})$, we adopt the Fourier-conjugate frequency space (ω space) and represent the dipole correlation kernel $K_p(\omega)$ and the electrostatic field $\mathbf{E} = \mathbf{E}(\mathbf{r})$ in this space. In contrast with other treatments²³, we note that \mathbf{P} and \mathbf{E} are indeed proportional but the proportionality constant is ω dependent²⁴. Thus, in ω space, we get

$$F(\mathbf{P})(\omega) = K_p(\omega)F(\mathbf{E})(\omega), \quad (3)$$

where F denotes three-dimensional Fourier transform $F(\mathbf{f})(\omega) = (2\pi)^{-3/2} \int e^{i\omega \cdot \mathbf{r}} \mathbf{f}(\mathbf{r}) d\mathbf{r}$, and the kernel $K_p(\omega)$ is the Lorentzian $K_p(\omega) = (\epsilon_b - \epsilon_o)/(1 + (\tau(\mathbf{r})c)^2|\omega|^2)$, with $\tau(\mathbf{r})c$ = position-dependent dielectric relaxation scale $\approx 3 \text{ cm}$ for $\tau = \tau_b \approx 100 \text{ ps}$ (c = speed of light), ϵ_b = bulk permittivity and ϵ_o = vacuum permittivity. Because $\mathbf{P}(\mathbf{r})$ satisfies the Debye relation $\nabla \cdot (\epsilon_o \mathbf{E} + \mathbf{P})(\mathbf{r}) = \rho(\mathbf{r})$, where $\rho(\mathbf{r})$ = charge density, equation (4) yields the following equation in \mathbf{r} space²⁵:

$$\nabla \cdot [F^{-1}(K)(\mathbf{r} - \mathbf{r}')\mathbf{E}(\mathbf{r}') d\mathbf{r}'] = \rho(\mathbf{r}), \quad (4)$$

with $K(\omega) = \epsilon_o + K_p(\omega)$. The convolution $[F^{-1}(K)(\mathbf{r} - \mathbf{r}')\mathbf{E}(\mathbf{r}') d\mathbf{r}']$ captures the correlation of the dipoles with the electrostatic field. Note that equation (4) is not the Poisson-Boltzmann equation, which requires a proportionality between the fields \mathbf{E} and \mathbf{P} under the *ad hoc* assumption $K(\omega) = \text{constant}$.

Upon water confinement, the dielectric relaxation undergoes a frequency red-shift arising from the reduction in hydrogen-bond partnerships that translates to a reduction in dipole orientation possibilities. Thus, at position \mathbf{r} , the relaxation time is $\tau = \tau_b \exp(B(g(\mathbf{r}))/k_B T)$, where the kinetic barrier $B(g(\mathbf{r})) = -k_B T \ln(g(\mathbf{r})/4)$ yields $\tau(\mathbf{r}) = \tau_b (g(\mathbf{r})/4)^{-1}$. Thus, for charge distribution,

$$\rho(\mathbf{r}) = \sum_m \in L 4\pi q_m \delta(\mathbf{r} - \mathbf{r}_m), \quad (5)$$

with L = set of charges on the protein surface labelled by index m , the g -dependent polarization is obtained from equation (4) (Supplementary Information):

$$\begin{aligned} \mathbf{P}(\mathbf{r}) &= [F^{-1}(K_p)(\mathbf{r} - \mathbf{r}')\mathbf{E}(\mathbf{r}') d\mathbf{r}'] \\ &= (2\pi)^{-3} \sum_m \in L [d\mathbf{r}' F^{-1}(K_p)(\mathbf{r} - \mathbf{r}') \nabla_{\mathbf{r}'} \cdot d\mathbf{\omega} e^{-i\omega \cdot (\mathbf{r}' - \mathbf{r}_m)} 4\pi q_m / [|\omega|^2 K(\omega)]]. \end{aligned} \quad (6)$$

Spatially dependent coordination $g = g(\mathbf{r})$. The time-averaged scalar field $g = g(\mathbf{r})$ was obtained from classical trajectories generated by molecular dynamics. The computations started with the PDB structure of a free (uncomplexed) protein molecule embedded in a pre-equilibrated cell of explicitly represented water molecules and counterions^{18,19}. The molecular-dynamics trajectories were generated by adopting an integration time step of 2 fs in an NPT ensemble with box size 10^3 nm^3 and periodic boundary conditions²⁶. The box size was calibrated so that the solvation shell extended at least 10 Å from the protein surface at all times. The long-range electrostatics were treated using the particle mesh Ewald summation method²⁷. A Nosé-Hoover thermostat²⁸ was used to maintain the temperature at 300 K, and a Tip3P water model with the optimized potential for liquid simulations (OPLS) force field was adopted^{18,19}. A barostat scheme was maintained through a dedicated routine with the pressure held constant at 1 atm. using a weak-coupling algorithm²⁹. After equilibration for 300 ns, g values averaged over a time span of 100 ns were determined for each point in space.

PWIT as promoter of protein-protein associations. The PWIT computed using equations (2) and (6) is generated by interfacial hotspots of red-shifted dielectric relaxation ($g(\mathbf{r}) < 4$, $\tau(\mathbf{r}) > \tau_b$). The most common spots involve hindered polar hydration generated by SABHBs (Fig. 1a). Taken collectively, the SABHBs contribute $73 \pm 5\%$ to the interfacial tension (Supplementary Information). The results are validated by showing that the inferred patches of interfacial tension

promote protein associations, a conclusion supported by the tight correlation ($r^2 = 0.83$) between the total area of surface patches begetting PWIT (increasing the value of the integral in equation (2)) in free complex subunits, and the total protein-protein interfacial area of protein complexes (Supplementary Fig. 2a). The relevance of PWIT as a molecular determinant of protein-protein interactions is further validated by showing that inferred tension patches actually coincide with hotspots at complex interfaces experimentally identified by mutational scanning (Supplementary Fig. 2b, c).

Identification of SABHBs in soluble proteins. The extent of protection of a backbone hydrogen bond, ζ , was computed directly from PDB structural coordinates by determining the number of side-chain non-polar groups contained within a desolvation domain around the bond^{20,22}. This domain was defined as two intersecting spheres of fixed radius (approximate thickness of three water layers) centred at the α -carbons of the residues paired by the hydrogen bond. In structures of soluble proteins, backbone hydrogen bonds are protected on average by $\zeta = 26.6 \pm 7.5$ non-polar groups for a desolvation sphere of radius 6 Å. SABHBs lie in the tails of the distribution: that is, their microenvironment contains 19 or fewer non-polar groups ($\zeta \leq 19$), so their ζ value is below the mean minus one standard deviation.

Sequence-based identification of SABHBs. SABHBs represent structural vulnerabilities that have been characterized as belonging to a twilight zone between order and native disorder. This characterization is justified by a strong correlation between intramolecular hydrogen-bond protection, ζ , and propensity for structural disorder (f_d) (Supplementary Fig. 5). The correlation reveals that the inability to exclude water intramolecularly from pre-formed hydrogen bonds is causative of the loss of structural integrity. The disorder propensity is accurately quantified by a sequence-based score generated by the program PONDR-VLXT³⁰, a predictor of native disorder that takes into account residue attributes such as hydrophilicity, aromaticity and their distribution within the window interrogated. The disorder score ($0 \leq f_d \leq 1$) is assigned to each residue within a sliding window, representing the predicted propensity of the residue to be in a disordered region ($f_d = 1$, certainty of disorder; $f_d = 0$, certainty of order). Only 6% of 1,100 non-homologous PDB proteins gave false-positive predictions of disorder in sequence windows of 40 amino acids^{22,30}. The strong correlation (Supplementary Fig. 5) between the disorder score of a residue and extent of protection of the hydrogen bond engaging the residue (if any) provides a sequence-based method of inference of SABHBs and supports the picture that such bonds belong to an order-disorder twilight zone²². Thus SABHBs can be safely inferred in regions where the disorder score lies in the range $0.35 \leq f_d < 0.95$, which corresponds to a marginal BHB protection with $7 \leq \zeta \leq 19$ (Supplementary Fig. 5).

Evaluation of homology models. The homology models based on template PDB structures from orthologous proteins were evaluated, ranked and ultimately selected using ProSA¹⁵, based on the minimization of $(Z_{\text{mod}} - Z_{\text{temp}})/Z_{\text{temp}}$, where Z_{mod} and Z_{temp} are the Z scores of model and template. The Z score of a structure or template-based model is the energetic gap between the structure and an average over an ensemble of random conformations for the protein chain¹⁵.

23. Schutz, C. N. & Warshel, A. What are the dielectric constants of proteins and how to validate electrostatic models? *Proteins Struct. Funct. Genet.* **44**, 400–417 (2001).
24. Scott, R., Boland, M., Rogale, K. & Fernández, A. Continuum equations for dielectric response to macromolecular assemblies at the nanoscale. *J. Phys. A* **37**, 9791–9803 (2004).
25. Fernández, A., Sosnick, T. R. & Colubri, A. Dynamics of hydrogen-bond desolvation in folding proteins. *J. Mol. Biol.* **321**, 659–675 (2002).
26. Lindahl, E., Hess, B. & Van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7**, 302–317 (2001).
27. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
28. Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
29. Berendsen, H. J., Postma, J. P., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
30. Li, X., Romero, P., Rani, M., Dunker, A. K. & Obradovic, Z. Predicting protein disorder for N-, C-, and internal regions. *Genome Inform.* **10**, 30–40 (1999).

Control of visual cortical signals by prefrontal dopamine

Behrad Noudoost¹ & Tirin Moore¹

The prefrontal cortex is thought to modulate sensory signals in posterior cortices during top-down attention^{1,2}, but little is known about the underlying neural circuitry. Experimental and clinical evidence indicate that prefrontal dopamine has an important role in cognitive functions³, acting predominantly through D1 receptors. Here we show that dopamine D1 receptors mediate prefrontal control of signals in the visual cortex of macaques (*Macaca mulatta*). We pharmacologically altered D1-receptor-mediated activity in the frontal eye field of the prefrontal cortex and measured the effect on the responses of neurons in area V4 of the visual cortex. This manipulation was sufficient to enhance the magnitude, the orientation selectivity and the reliability of V4 visual responses to an extent comparable with the known effects of top-down attention. The enhancement of V4 signals was restricted to neurons with response fields overlapping the part of visual space affected by the D1 receptor manipulation. Altering either D1- or D2-receptor-mediated frontal eye field activity increased saccadic target selection but the D2 receptor manipulation did not enhance V4 signals. Our results identify a role for D1 receptors in mediating the control of visual cortical signals by the prefrontal cortex and suggest how processing in sensory areas could be altered in mental disorders involving prefrontal dopamine.

Dopamine D1 receptors (D1Rs) are expressed by about one-quarter of all neurons in the prefrontal cortex and are localized primarily in superficial and deep layers^{4–6}. Microiontophoretic application of the selective D1R antagonist SCH23390⁷ at certain doses can increase the persistent, working-memory-related component of single-neuron activity in the dorsolateral prefrontal cortex^{3,8,9}. Given the role of the prefrontal cortex in visual attention^{1,2}, we hypothesized that D1Rs might also mediate the top-down control of visual cortical signals by the prefrontal cortex. If so, then changes in D1R-mediated prefrontal cortex activity might be sufficient to modulate signals in the posterior visual cortex, similar to the modulation observed during selective attention¹⁰. The prefrontal cortex's influence on the visual cortex is achieved in part by the frontal eye field (FEF)^{11,12}, an oculomotor area within the posterior prefrontal cortex. The FEF has a well-established role in saccadic target selection¹³, but recent evidence also implicates this area in the control of spatial attention^{2,14,15}. To test our hypothesis, we locally infused¹⁶ small volumes (0.5–1 μ l) of SCH23390 into sites in the FEF of macaques performing fixation and eye movement tasks (Fig. 1a, b and Supplementary Fig. 1). We measured the effects of the FEF infusion on target selection using a free-choice saccade task¹⁷. In this task, monkeys were rewarded for choosing between two saccadic targets, one located within the FEF response field and one in the opposite hemifield. In the same experiment, we recorded the visual responses of single neurons in area V4 during fixation. In particular, we recorded neurons with response fields that overlapped the FEF response field. Thus, we tested the effects of the D1R manipulation on both visual cortical signals and saccadic target selection.

We found that altering D1R-mediated activity at FEF sites increased the tendency of monkeys to choose targets appearing within the FEF response field (Fig. 1b). In the free-choice task, the temporal onset of

the two targets was systematically varied such that the FEF response field stimulus could appear earlier or later than the opposite stimulus. A monkey's tendency to select the FEF response field target could then

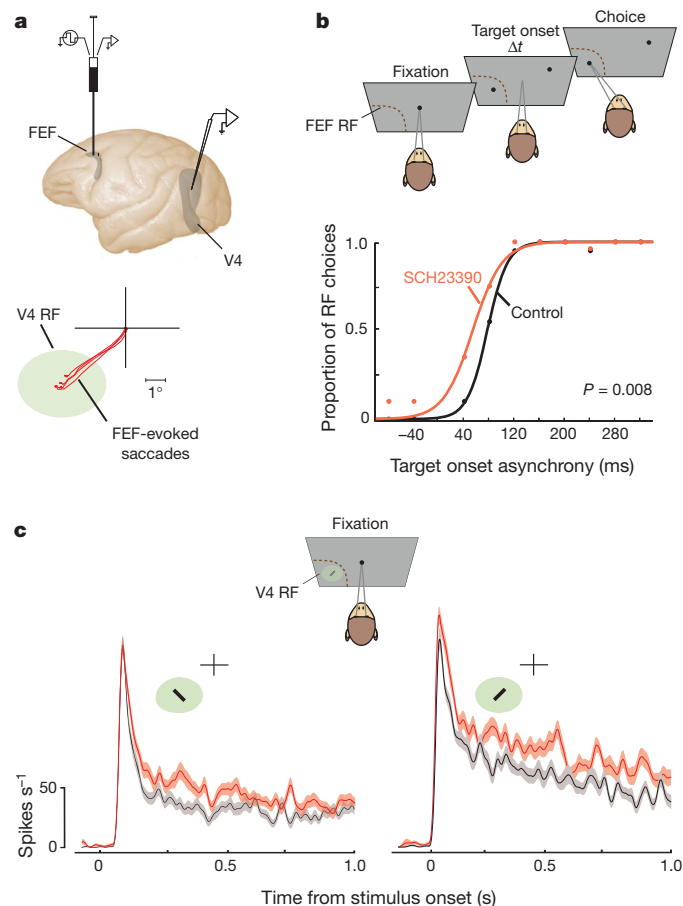


Figure 1 | Local manipulation of D1R-mediated activity in the FEF during single-neuron electrophysiology in area V4. **a**, Lateral view of the macaque brain depicting the location of a recording microsyringe in the FEF and of recording sites in area V4. Bottom diagram shows saccades evoked via electrical microstimulation at the infusion site (red traces) and the response field (RF, green ellipse) of a recorded V4 neuron in an example experiment. **b**, Double-target saccade task used to measure the monkey's tendency to make saccades to a target within the FEF response field versus one at an opposite location across varying temporal onset asynchronies. Positive asynchrony values denote earlier onset of FEF response field targets. Bottom plot shows the leftward shift in the PES, indicating more FEF response field choices, after infusion of SCH23390 into an FEF site. **c**, Visual responses of a V4 neuron with a response field that overlapped the FEF response field, measured during passive fixation. The plot shows mean \pm s.e.m. of visual responses to a bar stimulus presented at orthogonal orientations before (grey) and after (red) the infusion of SCH23390 at the FEF site.

¹Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Fairchild building, 299 Campus Drive West, Stanford, California 94305, USA.

be measured as the temporal onset asynchrony required for an equal probability of selecting either stimulus; we termed this the point of equal selection (PES). In the example experiment shown, the monkey chose the FEF response field target as often as the opposite target when the former appeared 76 ms earlier ($PES = 76$). However, infusion of SCH23390 (0.85 μ l) into the FEF reduced the PES by 23 ms (binary logistic regression, $P = 0.007$), thereby increasing the proportion of FEF response field target choices.

In the same experiment, we also measured the responses of V4 neurons to oriented bars during fixation in a separate task (Fig. 1c and Supplementary Methods). We found that the increase in target selection after the SCH23390 infusion was accompanied by an enhanced V4 neuronal response to oriented bars appearing within the overlapping V4 and FEF response fields. The example neuron shown was selective for orientation: it responded more to the 45° than to the 135° bar stimulus ($P < 10^{-3}$). After the infusion of SCH23390, there was a significant increase in the overall visual response of this neuron as well as a significant increase in the differential response to the two orientations (two-way analysis of variance, SCH23390 effect, $P < 10^{-3}$; SCH23390–orientation interaction, $P < 10^{-3}$). Thus, the local perturbation of D1R-mediated FEF activity not only caused the monkey to select FEF response field stimuli as saccade targets more frequently, it also led to enhanced and more selective visual responses of a V4 neuron representing the same part of space.

We studied the visual responses of 37 V4 neurons with response fields that overlapped the response fields of FEF infusion sites. The average (mean \pm s.e.m.) distance between V4 response field and FEF response field centres was 0.71 ± 0.07 degrees of visual angle (d.v.a.) (Fig. 2a). As with the example neuron, we measured the responses of all neurons to oriented bars appearing in their response field during a 1 s fixation period (Fig. 2b). Before the onset of the visual stimulus, there was a significant elevation in baseline activity after the D1R manipulation (Δ baseline = 0.077 ± 0.186 , $P = 0.030$). In addition to the baseline increase, the visually driven response of V4 neurons was enhanced by 17% above the control response (Δ response = 0.121 ± 0.054 , $P = 0.018$). We confirmed that the enhancement in the visual response was not due to systematic changes in eye position during stimulus presentation (Supplementary Fig. 2). The enhancement of the visual response was independently significant for both preferred (Δ preferred = 0.264 ± 0.087 ; $P = 0.004$) and non-preferred stimuli (Δ non-preferred = 0.132 ± 0.062 ; $P = 0.032$). There was also an increase in the response difference between the preferred and non-preferred orientations (Δ response difference = 0.132 ± 0.041 ; $P = 0.004$) (Supplementary Fig. 3), indicating an increase in orientation selectivity. To measure selectivity more quantitatively, we used a receiver-operating characteristic (ROC) analysis to quantify the degree to which each neuron's responses could be used to judge stimulus orientation (Fig. 2c). This analysis confirmed that V4 neurons were more orientation selective after changes in D1R-mediated FEF activity (Δ ROC area = 0.035 ± 0.009 , $P < 10^{-3}$). The enhancement in the magnitude and selectivity of the V4 response was accompanied by a decrease in the trial-to-trial variability of visual responses. We measured the variability of V4 responses across trials by computing the Fano factor, which is the variance in the spike count divided by its mean. We found that the Fano factor of V4 responses was reduced after the D1R manipulation (Δ FF = -0.105 ± 0.045 ; $P < 10^{-3}$) (Fig. 2d and Supplementary Fig. 4). All three V4 effects were comparable in magnitude to the known effects of top-down attention and consistent with a multiplicative increase in the gain of visual signals^{18,19} (Fig. 2e).

The effect of the D1R manipulation on saccadic target selection was highly consistent across the two monkeys tested. In 21 double-target experiments, the PES was reduced in every case (Fig. 3a). The mean PES shifted in favour of the FEF response field stimulus by an average of 27 ms (Δ PES = -26.934 ± 3.086 , $P < 10^{-3}$), significantly increasing the overall proportion of FEF response field choices

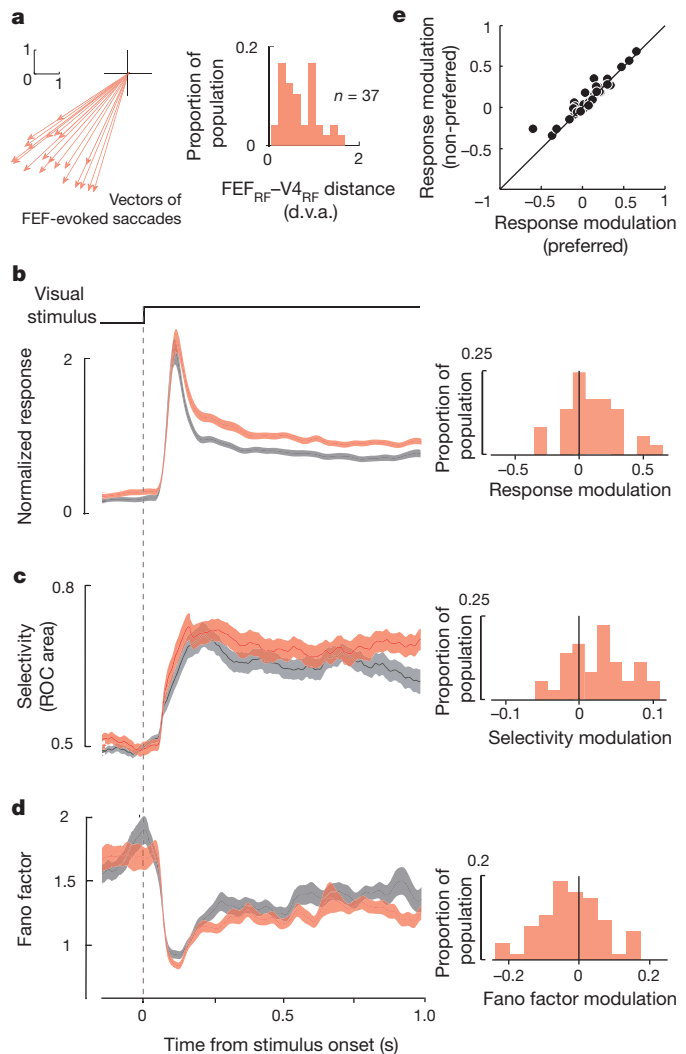


Figure 2 | Manipulation of D1R-mediated activity enhances V4 visual signals. **a**, Average vectors of saccades evoked at all FEF sites that overlapped V4 response fields (left panel). The distribution of distances between the endpoints of evoked saccades and the centres of overlapping V4 response fields for 37 V4 neurons is shown in the right panel. **b–d**, The mean normalized response magnitude (**b**), orientation selectivity (**c**) and response variability (Fano factor) (**d**) of V4 neurons before (grey) and after (red) microinfusion of SCH23390 into the FEF. Means \pm s.e.m. are shown within a 100-ms moving window measured during the 1-s response field stimulus presentation (top event plot). Histograms to the right of each response profile show the distributions of modulation indices for response magnitude (**b**), selectivity (**c**) and variability (**d**) across the population of neurons. **e**, Comparison of V4 response modulation after the SCH23390 infusion for preferred and non-preferred response field stimuli.

(chi-squared = 80.60, $P < 10^{-3}$) and thus indicating that the D1R manipulation increased the monkeys' tendency to target FEF response field stimuli. The increase in target selection was apparent across a range of drug dosages (Supplementary Fig. 5). In addition to the D1R manipulation, we tested the effects of the D2R agonist quinpirole. Previous studies using this drug found that it does not affect persistent activity but rather increases saccade-related activity within the dorso-lateral prefrontal cortex²⁰. We found that local manipulation of D2R-mediated FEF activity, like the D1R manipulation, increased the selection of FEF response field targets (Fig. 3a). The PES shifted by an average of 22 ms (Δ PES = -21.993 ± 6.758 , $P = 0.010$), increasing the proportion of FEF response field choices (chi-squared = 13.86, $P < 10^{-3}$). Thus, the D1R- and D2R-mediated manipulations of FEF activity resulted in equivalent increases in saccadic target selection.

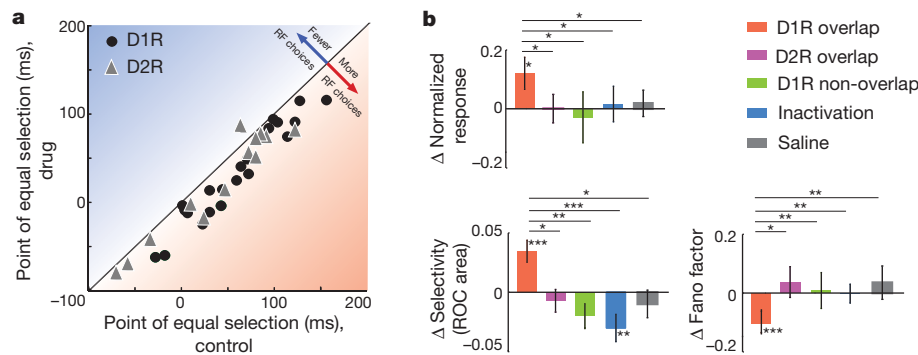


Figure 3 | Changes in saccadic target selection and V4 visual responses. **a**, Scatter plot shows the consistent increase in FEF response field target choices (decrease in PES) after manipulation of both D1R-mediated (circles) and D2R-mediated (triangles) FEF activity. For both drug effects, the increase in FEF response field target selection was constant across a range of control PES values; the slope in the linear fit did not differ significantly from unity in either case

Despite the increase in target selection, manipulation of D2R-mediated activity in the FEF failed to enhance the responses of V4 neurons. We found no significant effect on the visual response magnitude, orientation selectivity or response variability of V4 neurons after the D2R manipulation (Δ response = 0.001 ± 0.048 , $P = 0.999$; Δ ROC area = -0.007 ± 0.010 , $P = 0.426$; Δ FF = 0.037 ± 0.052 , $P = 0.338$; $n = 15$) (Fig. 3b). Moreover, the changes in these measures were all significantly different from the changes we observed after the D1R manipulation (Δ response_{D2R} < Δ response_{D1R}, $P = 0.045$; Δ selectivity_{D2R} < Δ selectivity_{D1R}, $P = 0.011$; Δ FF_{D2R} > Δ FF_{D1R}, $P = 0.019$). Thus, the equivalent effects of D1R and D2R manipulations on saccadic target selection were accompanied by contrasting effects in V4, with the enhancement of visual signals being specific to D1R-mediated activity. We also found that this enhancement was confined to V4 neurons with response fields that overlapped the FEF response field. For V4 neurons with response fields that did not overlap the FEF response field (mean distance between V4 response field and FEF response field = 9.00 ± 0.86 d.v.a.; $n = 15$), we found no significant effect of the D1R manipulation on response magnitude (Δ response = -0.028 ± 0.087 , $P = 0.9780$), orientation selectivity (Δ ROC area = -0.017 ± 0.010 , $P = 0.187$) or the Fano factor (Δ FF = 0.010 ± 0.043 , $P = 0.688$). Of note, the changes in these measures were all significantly different from the changes observed in neurons with overlapping response fields (Δ response_{non-overlap} < Δ response_{overlap}, $P = 0.044$; Δ selectivity_{non-overlap} < Δ selectivity_{overlap}, $P = 0.007$; Δ FF_{non-overlap} > Δ FF_{overlap}, $P = 0.034$) (Fig. 3b). Thus, the enhancement in visual cortical signalling produced by manipulation of D1R-mediated FEF activity was spatially specific.

We also tested the effect of complete inactivation of FEF sites on the responses of V4 neurons with overlapping response fields. Previous studies have shown that local inactivation of the FEF disrupts saccadic target selection and impairs attention^{17,21}. We therefore wondered if inactivation could reduce the components of V4 responses that were enhanced by the D1R manipulation. We locally inactivated FEF sites using the GABA_A (γ -aminobutyric acid subtype A) receptor agonist muscimol. Unlike the sparsely expressed D1Rs, GABA_A receptors are expressed by all neurons in all cortical layers²². As in previous studies, local inactivation of FEF sites with muscimol decreased the targeting of FEF response field stimuli. It also significantly reduced V4 orientation selectivity (Δ ROC area = -0.030 ± 0.011 , $P = 0.003$; $n = 33$). However, the inactivation did not change the response magnitude or variability of V4 neurons (Δ response = 0.016 ± 0.061 , $P = 0.809$; Δ FF = -0.002 ± 0.023 , $P = 0.921$) (Fig. 3b). Thus, in contrast to the D1R manipulation which altered all three components of V4 activity, complete inactivation altered only one. All three inactivation effects were significantly different from the D1R effects

(D1R: slope = 0.96, $P = 0.552$; D2R: slope = 0.97, $P = 0.502$). **b**, Changes in response magnitude, orientation selectivity and response variability (Fano factor) after each drug manipulation. Changes shown are mean differences from pre-infusion values. Error bars denote s.e.m.; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

(Δ response_{muscimol} < Δ response_{D1R}, $P = 0.024$; Δ selectivity_{muscimol} < Δ selectivity_{D1R}, $P < 10^{-3}$; Δ FF_{muscimol} > Δ FF_{D1R}, $P = 0.007$). Although the reduction in orientation selectivity is consistent with previous electrical microstimulation studies¹² and with the effects of inactivation on orientation discrimination²¹, the lack of a reduction in response magnitude may seem inconsistent. However, we suggest that this difference is due to variation between experimental paradigms (Supplementary Discussion). Finally, we tested for any effect of vehicle (saline) infusion into the FEF. The infusion of saline failed to change the response magnitude, selectivity or variability of V4 neurons (Δ response = 0.018 ± 0.048 , $P = 0.380$; Δ ROC area = -0.010 ± 0.013 , $P = 0.569$; Δ FF = -0.035 ± 0.061 , $P = 0.179$; $n = 12$) (Fig. 3b). All three measures were significantly different from the D1R effects (Δ response_{saline} < Δ response_{D1R}, $P = 0.045$; Δ selectivity_{saline} < Δ selectivity_{D1R}, $P = 0.013$; Δ FF_{saline} > Δ FF_{D1R}, $P = 0.009$).

Our results identify prefrontal D1Rs as a component of the neural circuitry controlling signals in the visual cortex. Manipulation of D1R-mediated FEF activity was sufficient to enhance the magnitude, reliability and visual selectivity of neuronal responses in area V4, three known effects of visual attention. The observed enhancement might account for the benefits in visually guided behaviour that accompany attentional deployment (Supplementary Fig. 6), although a causal link between attentional modulation of visual cortical signals and visual perception remains to be established. We have demonstrated that visual representations in posterior areas can be altered merely by changes in dopamine tone in the prefrontal cortex. Given the complex effects of dopamine through D1Rs, one might predict that at 'optimum' dopamine levels⁹, optimal top-down control of visual cortical signals would be achieved.

The circuitry underlying top-down control of the visual cortex probably involves several different neuromodulators²³ and an array of different brain structures²⁴. Our results show that this circuitry involves prefrontal dopamine acting via D1Rs. In the dorsolateral prefrontal cortex, dopamine D1Rs are thought to modulate recurrent glutamatergic connections, thereby influencing activity related to working memory in this area^{25,26}. This study shows that D1Rs contribute to the FEF's control of visual signals by an analogous mechanism, namely by modulating long-range, recurrent connections between the FEF and the visual cortex (Supplementary Fig. 7). Because FEF neurons in the superficial layer are reciprocally connected with neurons in V4^{2,27}, dopaminergic modulation of these connections via D1Rs in the superficial layer would be expected to mediate the FEF's control of V4 signals. The specificity of V4 effects to D1Rs, rather than D2Rs, might be explained by the relative absence of D2Rs in superficial layers of the prefrontal cortex⁴⁻⁶. The equivalent effects of D1R and D2R

manipulations on target selection might be explained by the presence of both receptor subtypes in infragranular layers of the cortex^{4–6}, where layer-V FEF neurons project to the superior colliculus²⁷.

Impairments in saccadic control are prominent among the impairments exhibited in attention deficit/hyperactivity disorder (ADHD)²⁸. The observed influence of prefrontal DIRs on saccadic target selection and visual cortical signals, combined with their known influence on persistent activity, may explain the behavioural links between saccadic control, attention and working memory²⁹ and the coincidence of their corresponding impairments in ADHD³⁰.

METHODS SUMMARY

The effects of pharmacological perturbation of FEF activity on target selection and the visual responses of V4 neurons were studied in three macaques (*Macaca mulatta*) performing fixation and eye movement tasks (Supplementary Methods). All experimental procedures were in accordance with the National Institutes of Health guide for the care and use of laboratory animals and with the Society for Neuroscience guidelines and policies. They were also approved by the Stanford University animal care and use committee. Eye position was monitored with a scleral search coil. In each experiment, we infused small volumes of drug into sites in the FEF through a surgically implanted titanium chamber overlying the arcuate sulcus using a custom-made recording microinjector. We identified FEF sites by eliciting short-latency, fixed-vector saccadic eye movements with trains (50–100 ms) of biphasic current pulses ($\leq 50 \mu\text{A}$; 250 Hz; 0.25 ms duration). In the same experiment, recordings from V4 neurons were made through a chamber overlying the prelunate gyrus. Response fields of V4 neurons were all located in the lower quadrant of the contralateral hemifield ($<12^\circ$ eccentricity). The position of the FEF microinjector was adjusted so that the saccade elicited by FEF microstimulation shifted the monkey's gaze either to within the V4 response field (overlapping) or far outside it (non-overlapping).

Received 20 May 2010; accepted 14 March 2011.

Published online 15 May 2011.

- Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
- Noudoost, B., Chang, M. H., Steinmetz, N. A. & Moore, T. Top-down control of visual attention. *Curr. Opin. Neurobiol.* **20**, 183–190 (2010).
- Robbins, T. W. & Arnsten, A. F. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* **32**, 267–287 (2009).
- Lidow, M. S., Goldman-Rakic, P. S., Gallager, D. W. & Rakic, P. Distribution of dopaminergic receptors in the primate cerebral cortex: quantitative autoradiographic analysis using [³H]raclopride, [³H]spiperone and [³H]SCH23390. *Neuroscience* **40**, 657–671 (1991).
- Lidow, M. S., Wang, F., Cao, Y. & Goldman-Rakic, P. S. Layer V neurons bear the majority of mRNAs encoding the five distinct dopamine receptor subtypes in the primate prefrontal cortex. *Synapse* **28**, 10–20 (1998).
- Santana, N., Mengod, G. & Artigas, F. Quantitative analysis of the expression of dopamine D1 and D2 receptors in pyramidal and GABAergic neurons of the rat prefrontal cortex. *Cereb. Cortex* **19**, 849–860 (2009).
- Bourne, J. A. SCH 23390: the first selective dopamine D1-like receptor antagonist. *CNS Drug Rev.* **7**, 399–414 (2001).
- Williams, G. V. & Goldman-Rakic, P. S. Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* **376**, 572–575 (1995).
- Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V. & Arnsten, A. F. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nature Neurosci.* **10**, 376–384 (2007).
- Reynolds, J. H. & Chelazzi, L. Attentional modulation of visual processing. *Annu. Rev. Neurosci.* **27**, 611–647 (2004).
- Gregoriou, G. G., Gotts, S. J., Zhou, H. & Desimone, R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* **324**, 1207–1210 (2009).
- Moore, T. & Armstrong, K. M. Selective gating of visual signals by microstimulation of frontal cortex. *Nature* **421**, 370–373 (2003).
- Schall, J. D. & Hanes, D. P. Neural basis of saccade target selection in frontal eye field during visual search. *Nature* **366**, 467–469 (1993).
- Thompson, K. G., Biscoe, K. L. & Sato, T. R. Neuronal basis of covert spatial attention in the frontal eye field. *J. Neurosci.* **25**, 9479–9487 (2005).
- Armstrong, K. M., Chang, M. H. & Moore, T. Selection and maintenance of spatial information by frontal eye field neurons. *J. Neurosci.* **29**, 15621–15629 (2009).
- Noudoost, B. & Moore, T. A reliable microinjector system for use in behaving monkeys. *J. Neurosci. Methods* **194**, 218–223 (2011).
- Schiller, P. H. & Tehovnik, E. J. Cortical inhibitory circuits in eye-movement generation. *Eur. J. Neurosci.* **18**, 3127–3133 (2003).
- Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* **55**, 131–141 (2007).
- McAdams, C. J. & Maunsell, J. H. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* **19**, 431–441 (1999).
- Wang, M., Vijayraghavan, S. & Goldman-Rakic, P. S. Selective D2 receptor actions on the functional circuitry of working memory. *Science* **303**, 853–856 (2004).
- Monosov, I. E. & Thompson, K. G. Frontal eye field activity enhances object identification during covert visual search. *J. Neurophysiol.* **102**, 3656–3672 (2009).
- Huntsman, M. M., Isackson, P. J. & Jones, E. G. Lamina-specific expression and activity-dependent regulation of seven GABA_A receptor subunit mRNAs in monkey visual cortex. *J. Neurosci.* **14**, 2236–2259 (1994).
- Herrero, J. L. et al. Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. *Nature* **454**, 1110–1114 (2008).
- McAlonan, K., Cavanaugh, J. & Wurtz, R. H. Guarding the gateway to cortex with attention in visual thalamus. *Nature* **456**, 391–394 (2008).
- Goldman-Rakic, P. S. Cellular basis of working memory. *Neuron* **14**, 477–485 (1995).
- Durstewitz, D., Seamans, J. K. & Sejnowski, T. J. Neurocomputational models of working memory. *Nature Neurosci.* **3** (Suppl.), 1184–1191 (2000).
- Pouget, P. et al. Visual and motor connectivity and the distribution of calcium-binding proteins in macaque frontal eye field: implications for saccade target selection. *Front. Neuroanat.* **3**, 1–14 (2009).
- Munoz, D. P., Armstrong, I. T., Hampton, K. A. & Moore, K. D. Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. *J. Neurophysiol.* **90**, 503–514 (2003).
- Awh, E., Armstrong, K. M. & Moore, T. Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends Cogn. Sci.* **10**, 124–130 (2006).
- Castellanos, F. X. & Tannock, R. Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nature Rev. Neurosci.* **3**, 617–628 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. S. Aldrich for technical assistance, N. Steinmetz for help with Fano factor analysis and W. T. Newsome, E. I. Knudsen, K. M. Armstrong and R. F. Squire for comments on the manuscript. This work was supported by NIH EY014924, NSF IOB-0546891, The McKnight Foundation and an IBRO Fellowship to B.N.

Author Contributions B.N. designed and performed experiments, analysed data and wrote the paper; T.M. designed and performed experiments and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.N. (behrad@stanford.edu).

COP1 is a tumour suppressor that causes degradation of ETS transcription factors

Alberto C. Vitari^{1†}, Kevin G. Leong², Kim Newton¹, Cindy Yee³, Karen O'Rourke¹, Jinfeng Liu⁴, Lilian Phu⁵, Rajesh Vij⁶, Ronald Ferrando⁷, Suzana S. Couto⁷, Sankar Mohan⁸, Ajay Pandita⁸, Jo-Anne Hongo⁶, David Arnott⁵, Ingrid E. Wertz⁹, Wei-Qiang Gao^{3†}, Dorothy M. French⁷ & Vishva M. Dixit¹

The proto-oncogenes *ETV1*, *ETV4* and *ETV5* encode transcription factors in the E26 transformation-specific (ETS) family, which includes the most frequently rearranged and overexpressed genes in prostate cancer^{1–4}. Despite being critical regulators of development, little is known about their post-translational regulation. Here we identify the ubiquitin ligase COP1 (also known as RFWD2) as a tumour suppressor that negatively regulates *ETV1*, *ETV4* and *ETV5*. *ETV1*, which is mutated in prostate cancer more often, was degraded after being ubiquitinated by COP1. Truncated *ETV1* encoded by prostate cancer translocation *TMPRSS2:ETV1* lacks the critical COP1 binding motifs and was 50-fold more stable than wild-type *ETV1*. Almost all patient translocations render *ETV1* insensitive to COP1, implying that this confers a selective advantage to prostate epithelial cells. Indeed, COP1 deficiency in mouse prostate elevated *ETV1* and produced increased cell proliferation, hyperplasia, and early prostate intraepithelial neoplasia. Combined loss of COP1 and PTEN enhanced the invasiveness of mouse prostate adenocarcinomas. Finally, rare human prostate cancer samples showed hemizygous loss of the *COP1* gene, loss of COP1 protein, and elevated *ETV1* protein while lacking a translocation event. These findings identify COP1 as a tumour suppressor whose downregulation promotes prostatic epithelial cell proliferation and tumorigenesis.

Mass spectrometry showed that *ETV1*, *ETV4*, and *ETV5* co-immunoprecipitated specifically with Flag-tagged COP1 from a mouse kidney epithelial cell line (Supplementary Fig. 1a). Known COP1-interacting proteins DET1 (ref. 5) and TRIB3 (ref. 6) also were co-immunoprecipitated (Supplementary Table 1). The *ETVs* each contain three potential COP1-binding motifs, with endogenous COP1 and *ETV1* interacting in LNCaP prostate cancer cells (Supplementary Fig. 1b–d). An inverse correlation between *ETV1* and COP1 proteins in prostate cancer cell lines suggested *ETV1* might be a COP1 substrate (Fig. 1a). For example, PC3 cells and their derivatives showed COP1 loss (data not shown), lacked detectable COP1, but did express *ETV1*. By contrast, COP1-expressing BPH1, BPH1025 and LNCaP cells lacked *ETV1* protein, despite LNCaP cells containing more *ETV1* mRNA than PC3 cells (Fig. 1a)⁷. Consistent with the notion that COP1 in LNCaP cells rendered newly synthesized *ETV1* unstable, *ETV1* protein, but not mRNA, was increased in LNCaP cells by either proteasome inhibition with MG-132 or Bortezomib (Fig. 1b) or siRNA knockdown of endogenous *COP1* (Fig. 1c). The latter increased the half-life of *ETV1* approximately 50-fold (Supplementary Fig. 2a). MG-132 did not cause *ETV1* accumulation in COP1-deficient PC3 cells (Supplementary Fig. 2b), consistent with *ETV1* being less subject to proteasomal degradation in the absence of COP1.

Mutation of the COP1 RING domain (C136A/C139A), which destroys its E3 ubiquitin ligase activity^{8,9}, or deletion of COP1 residues required for interaction with ligase component DET1 (COP1Δ24)⁵ prevented overexpressed COP1 from decreasing endogenous *ETV1* in PC3 cells (Supplementary Fig. 2c, d). MG-132, Bortezomib, and *DET1* knockdown in PC3 cells also abrogated *ETV1* destabilization by COP1 (Supplementary Fig. 2e, f). *ETV1* degradation by COP1 and DET1 was not limited to prostate cancer cells because co-expressed COP1 and DET1 reduced *ETV1* in DET1-deficient HCC1806 breast cancer cells (Supplementary Fig. 2g–i). The concerted action of COP1 and DET1 seen in *Arabidopsis thaliana*¹⁰ therefore is conserved in the regulation of mammalian *ETV1*.

Consistent with COP1 ubiquitin ligase activity being critical for *ETV1* degradation, *ETV1* co-expressed with COP1 and DET1 in 293T cells was ubiquitinated (Fig. 1d) and the polyubiquitin chains contained degradative K48 linkages¹¹ (Supplementary Fig. 2j). Conversely, *COP1* knockdown in LNCaP cells decreased ubiquitination of endogenous *ETV1* (Supplementary Fig. 2k). We predict that *ETV4* and *ETV5* are regulated similarly because of their conserved COP1 binding motifs and reduced expression in PC3 cells upon coexpression with COP1 and DET1 (Fig. 1e). Our findings support previous observations that COP1 regulates the stability of *ETV1*, *ETV4* and *ETV5* *in vitro*¹². Of note, the ETS transcription factor ERG, which lacks a COP1-binding motif, was not decreased by COP1 and DET1 (Fig. 1e).

Almost all reported *ETV1* chromosomal rearrangements in human prostate tumours yield N-terminally truncated *ETV1* (Δ*ETV1*), lacking the two N-terminal COP1 binding motifs (Supplementary Fig. 3a)^{1,2}. We proposed that these *ETV1* mutants evade COP1-mediated degradation and this contributes to their overexpression in prostate cancers. Unlike wild-type *ETV1*, Δ*ETV1* did not co-immunoprecipitate with RING mutant COP1 (Supplementary Fig. 3b, c). Nor was Δ*ETV1* polyubiquitinated and degraded by wild-type COP1 and DET1 (Supplementary Fig. 3d, e). To define the contribution of the three putative COP1-binding motifs or 'degrons' in *ETV1*, we mutated their conserved residues (VP to AA; Fig. 2a). Mutation of degreon 2 alone, but not degrons 1 or 3, decreased the interaction of *ETV1* with RING mutant COP1, and combining the degreon 1 and 2 mutations compromised the interaction further (Supplementary Fig. 3c). The degreon (1+2) mutant, like Δ*ETV1*, was no longer subject to COP1- and DET1-mediated degradation (Fig. 2b).

Unlike most oncogene products, *ETV1* overexpressed in cultured prostate cells promotes invasive behaviour rather than cell proliferation^{7,13}. COP1 silencing in LNCaP cells or overexpression in PC3 cells also had no effect on cell proliferation (data not shown). Instead, COP1 expression reduced PC3 cell migration through a collagen-coated

¹Department of Physiological Chemistry, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ²Department of Research Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ³Department of Molecular Biology, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ⁴Department of Bioinformatics and Computational Biology, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ⁵Department of Protein Chemistry, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ⁶Department of Antibody Engineering, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ⁷Department of Pathology, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ⁸Department of Molecular Diagnostics & Cancer Cell Biology, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. ⁹Department of Early Discovery Biochemistry, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. [†]Present addresses: Novartis Institutes for Biomedical Research, 4560 Horton Street, Emeryville, California 94608, USA (A.C.V.); Renji-MedX Clinical Stem Cell Research Center, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200127, China and Med-X Research Institute, Shanghai Jiao Tong University, Shanghai 200030, China (W.-Q.G.).

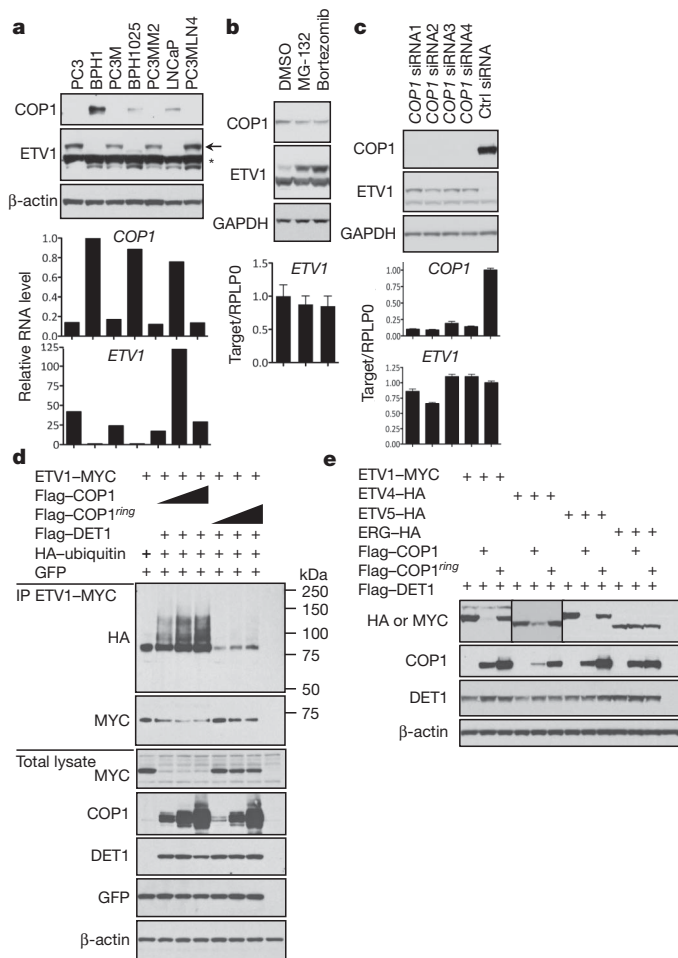


Figure 1 | The COP1/DET1 ubiquitin ligase regulates ETV1 turnover.

a, ETV1 expression in prostate cancer cell lines. Arrow, ETV1 protein; asterisk, bands of unknown identity. *COP1* and *ETV1* mRNA expression is plotted relative to the BPH1 sample, which is assigned a value of 1. **b**, **c**, ETV1 protein in LNCaP cells treated for 2 h with dimethylsulphoxide (DMSO), 20 μ M MG-132, or 2 μ M bortezomib (**b**) or transfected with four independent *COP1* siRNAs or a non-targeting (Ctrl) siRNA (**c**). Bars indicate the mean *ETV1* or *COP1* mRNA level \pm s.d. of triplicate wells after normalization to *RPLP0* gene expression. **d**, HEK293T cells were transfected and treated with 10 μ M MG-132 for 2 h before lysis. Solid triangles indicate increasing plasmid DNA. *COP1*^{ring} contained mutated residues C136A/C139A. Immunoprecipitations (IP) used SDS- and heat-denatured lysates. GFP, green fluorescent protein; HA, haemagglutinin. **e**, Transfected PC3 cells.

membrane, and this inhibition required the COP1 RING domain (Supplementary Fig. 3f). ETV1 degradation by COP1 likely contributed to decreased PC3 cell invasion because COP1 limited migration of cells coexpressing wild-type ETV1, but not the degron (1+2) ETV1 mutant that escapes COP1-mediated degradation (Fig. 2c).

Negative regulation of proto-oncogenic ETVs suggested a tumour suppressor role for COP1. In a renal graft model of prostate regeneration^{14,15}, *Cop1*^{+/+}, *Cop1*^{+/-} and *Cop1*^{-/-} prostate epithelial cells from gene-targeted mice (Supplementary Fig. 4) formed prostate structures at a similar frequency (Supplementary Table 2). Half of the *Cop1*^{-/-} structures, however, exhibited increased epithelial cell piling, loss of polarization, tortuous acini, and enhanced stromal cell invasion (Supplementary Fig. 5a, b). All structures expressed probasin, indicative of successful differentiation into prostatic secretory epithelium (Supplementary Fig. 5c), but COP1 loss was associated with increased expression of the luminal cell marker cytokeratin 18, as seen in prostate malignancies (Supplementary Fig. 5d, f). *Cop1*^{-/-} structures also showed a marked increase in Ki67-positive proliferating cells

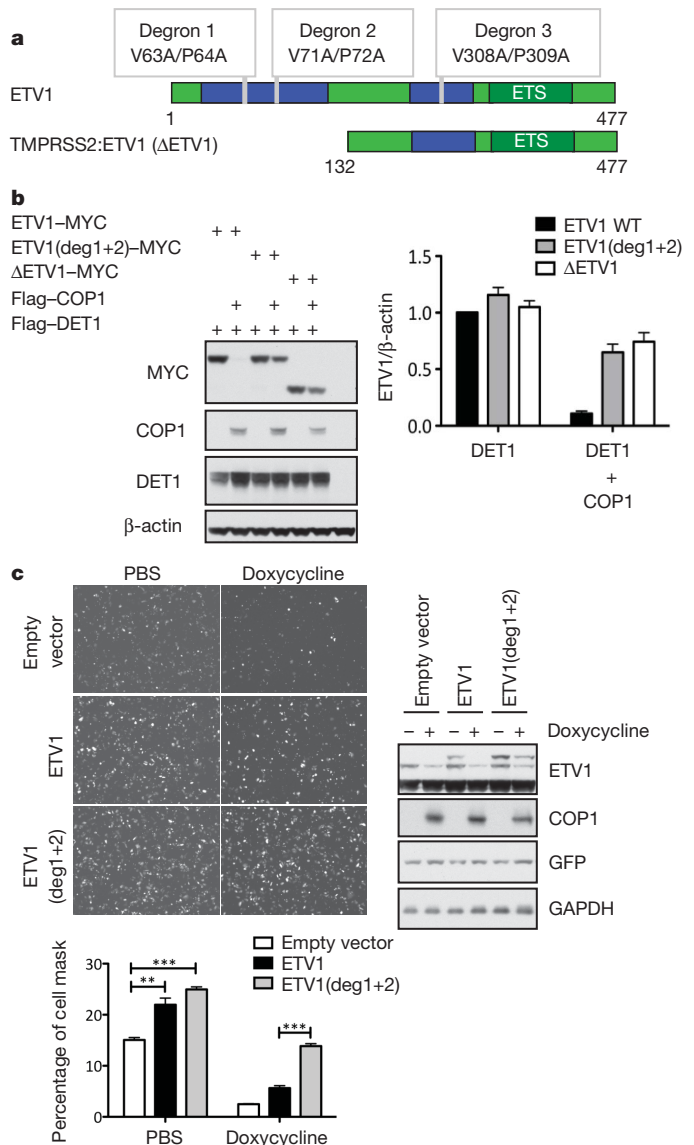
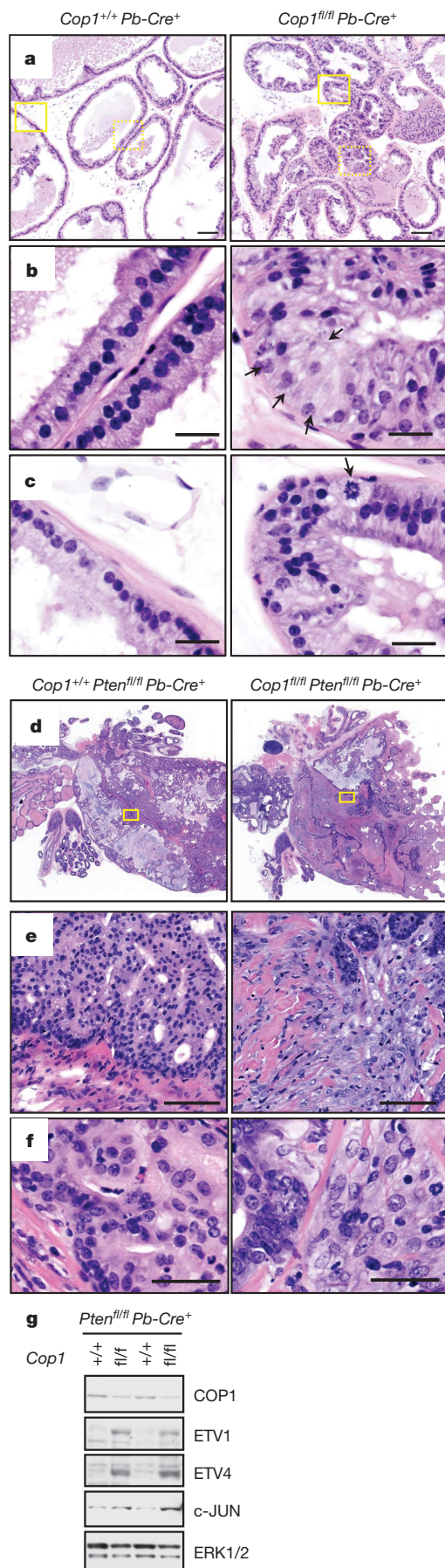


Figure 2 | Truncated ETV1 encoded by prostate cancer translocations is not degraded by the COP1/DET1 ubiquitin ligase.

a, Human ETV1 and truncated Δ ETV1 encoded by *TMPRSS2:ETV1*. COP1 binding motifs are indicated as degrons 1, 2, and 3. **b**, Transfected HEK293T cells. ETV1(deg1+2), ETV1(V63A/P64A/V71A/P72A). ETV1 protein was quantified and normalized to β -actin. Bars indicate the mean \pm s.d. of four replicates. **c**, PC3 cells stably expressing doxycycline-inducible COP1 were transfected with GFP plus empty vector, ETV1, or ETV1(deg1+2). Cells were tested for migration through a collagen-coated membrane after 40 h. GFP⁺ cells accessing the lower chamber were imaged and the GFP⁺ area (invading cells) plotted as percentage of cell mask. Bars show the mean \pm s.e.m. of four replicate wells. ***P* = 0.0025, ****P* < 0.0001, *t*-test.

(Supplementary Fig. 5e, g). These results indicate that COP1 deficiency causes aberrant prostatic epithelial cell growth.

Next we deleted *Cop1* in mouse prostatic epithelial cells *in vivo* with a probasin-Cre (*Pb-Cre*) transgene¹⁶. Wild-type (*Cop1*^{+/+} *Pb-Cre*⁺) and COP1-deficient (*Cop1*^{fl/fl} *Pb-Cre*⁺) prostates were comparable up to 24 weeks of age, but hyperplasia was evident in COP1-deficient prostate lobes by 40 weeks (Supplementary Table 3). By 52 weeks of age, hyperplasia was evident in all COP1-deficient prostates, and two out of six had developed low-grade mouse prostatic intraepithelial neoplasia (MPIN)¹⁷ in the ventral lobe (Fig. 3a–c and Supplementary Fig. 5h). MPIN lesions were confined to the gland and were reminiscent of abnormalities observed in transgenic mice expressing the *TMPRSS2:ETV1* gene product from the probasin promoter⁷. *Cop1* gene deletion correlated with decreased COP1 protein expression, increased cell proliferation, and



increased ETV1, ETV4 and c-JUN (Supplementary Fig. 5i–p). These data indicate that COP1 suppresses prostate tumour development and regulates the abundance of ETV1, ETV4 and c-JUN *in vivo*, all of which have been linked to prostate cancer^{1,2,18}.

Loss of the tumour suppressor PTEN is reported in ~50% of primary prostate cancers^{19,20}, and overexpression of ERG cooperates with PTEN loss in accelerating prostate cancer progression^{21,22}. We investigated whether COP1-deficiency in the prostate would cooperate similarly with PTEN loss and enhance tumour progression. By 30 weeks of age, prostates lacking COP1 and PTEN presented a more aggressive carcinoma when compared to prostates lacking only PTEN (Fig. 3d–f and Supplementary Table 3). Neoplastic epithelial cells in the *Cop1^{fl/fl} Pten^{fl/fl} Pb-Cre⁺* prostate extended beyond the basement membrane, invading the stromal compartment and the adjacent muscle bundles. Many cells were poorly differentiated with marked cellular atypia and nuclear pleomorphism (Fig. 3d–f). Loss of COP1 again correlated with elevated ETV1, ETV4 and c-JUN (Fig. 3g). Our observations suggest cooperation between COP1 and PTEN in suppressing tumorigenesis. We speculate that stabilization of ETV1 and ETV4 in the absence of COP1 upregulates matrix metalloproteinases and enhances cell invasion^{23,24}. Knockdown of endogenous *COP1* in LNCaP cells increased *MMP1*, *MMP7* and *MMP10* gene expression (Supplementary Fig. 6a–d), but this did not occur if both *COP1* and *ETV1* were knocked down. Indeed, microarray analyses showed that over 75% of the genes up-regulated in LNCaP cells after *COP1* knockdown were ETV1-dependent (Supplementary Fig. 6e). Silencing of COP1, ETV1 and c-JUN returned 92% of the upregulated genes to basal levels, indicating that ETV1 and c-JUN are both targets of COP1 suppression in LNCaP cells (Supplementary Fig. 6f, g and Supplementary Table 4).

Next we determined whether COP1 loss correlated with elevated ETV1 protein expression in human prostate cancer. Analysis of 166 comparative genomic hybridization (CGH) array data sets identified five cases with *COP1* loss (Fig. 4a). We retrieved three of these cases and focal loss of COP1 protein correlated with elevated ETV1 protein (Fig. 4b, c and data not shown). Regions of normal COP1 staining expressed minimal ETV1, supporting the inverse correlation between COP1 and ETV1 (Fig. 4d). Fluorescence *in situ* hybridization (FISH) in each case revealed loss of one copy of *COP1* (Fig. 4e and data not shown), and *in situ* hybridization indicated silenced *COP1* mRNA expression (Supplementary Fig. 7a). ETV1 and ERG break-apart FISH excluded that ETV1 overexpression in these samples was translocation-related (Fig. 4e and data not shown).

In parallel, we screened 120 human prostate cancer samples for elevated ETV1 protein expression by immunohistochemistry. Four cases exhibited focal overexpression of ETV1. One of these cases (HP2086) exhibited *ETV1* translocation, elevated *ETV1* mRNA and protein, and normal COP1 expression (Fig. 4e and data not shown). The remaining three cases lacked COP1 protein in areas staining strongly for ETV1 (Supplementary Fig. 7b–g) or c-JUN (Supplementary Fig. 8). FISH was successful for two of these three samples and there was loss of one copy of *COP1*. No *ETV1* or *ERG* translocations were detected (Supplementary Fig. 7h and data not shown). This scenario is probably analogous to what is observed with *PTEN* in prostate cancer: one *PTEN* allele is lost and the other is silenced²⁵. The mechanism(s) inactivating the remaining *COP1* allele is unknown. Regardless, these

Figure 3 | COP1 deficiency in prostatic epithelium causes hyperplasia, early MPIN, and enhances the effects of PTEN loss. **a–c**, Haematoxylin and eosin-stained ventral *Cop1^{+/+} Pb-Cre⁺* and *Cop1^{fl/fl} Pb-Cre⁺* prostates aged 52 weeks. Areas within dashed and solid boxes in **a** are magnified in **b** and **c**, respectively. Arrows indicate nuclear atypia. Scale bars: 100 μ m (**a**), 25 μ m (**b**, **c**). **d–f**, *Cop1^{+/+} Pten^{fl/fl} Pb-Cre⁺* and *Cop1^{fl/fl} Pten^{fl/fl} Pb-Cre⁺* prostates aged 30 weeks. Boxes in **d** are magnified in **e** and **f**, and contain characteristic MPIN (left panel) and invasive adenocarcinoma (right panel). Scale bars: 50 μ m (**e**), 25 μ m (**f**). **g**, ETV1 and ETV4 in *Cop1^{+/+} Pten^{fl/fl} Pb-Cre⁺* and *Cop1^{fl/fl} Pten^{fl/fl} Pb-Cre⁺* anterior prostates aged 30 weeks. ETV1 and ETV4 detected with antibodies 13G11 and 5F8, respectively.

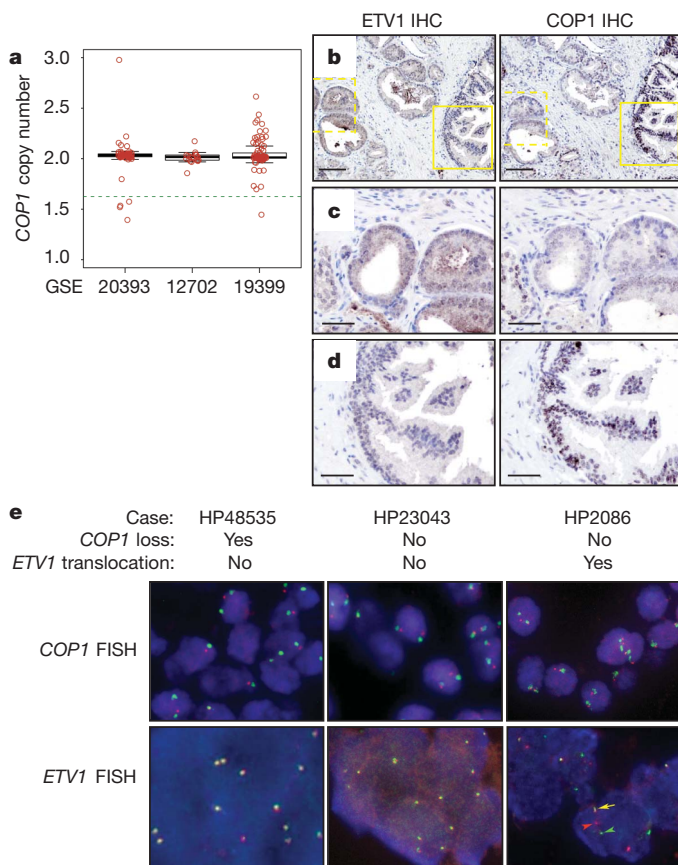


Figure 4 | Loss of *COP1* and decreased *ETV1* protein expression in human prostate adenocarcinomas. **a**, *COP1* DNA copy number in human prostate cancers. The dashed green line indicates the cutoff for copy number loss (\log_2 ratio ≤ -0.3). 5/166 tumours exhibit *COP1* loss. **b–d**, *ETV1* and *COP1* protein expression in *COP1*-deficient prostate adenocarcinoma case HP48535 detected by immunohistochemistry (IHC). Dashed or solid boxes in **b** are magnified in **c** and **d**, respectively. Scale bars: 100 μm (**b**); 50 μm (**c**, **d**). **e**, FISH analysis of HP48535 with *COP1* (red) and *CEP1* (green) probes (upper panels). Break-apart FISH assays for *ETV1* translocations (lower panels). A yellow fusion signal (arrow) is normal, whereas discrete red (arrowhead) and green (arrowhead) signals indicate *ETV1* translocation.

data support *COP1* deficiency being a mechanism for stabilization of substrates such as *ETV1* and *c-JUN* in human tumours. Even if human prostate cancer cells rarely lose *COP1*, they seem to have found an alternative means of evading negative regulation by *COP1*, namely loss of the *COP1* binding/degron motifs from labile proto-oncogene products such as *ETV1* (Supplementary Fig. 9). This finding may represent a general paradigm for oncogene fixation in transformed cells because labile oncoproteins have the advantage if they lose the degron responsible for their instability. For example, *c-MYC* overexpression in Burkitt's lymphoma is driven by translocation and mutation of residue T58, the latter being critical for its proteasome-mediated degradation^{26–29}. Our work, by extension, indicates that all translocations involving labile oncogenes are likely to eliminate or mutate the degron(s) that normally confer physiological instability.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 August 2010; accepted 17 March 2011.

Published online 15 May 2011.

1. Kumar-Sinha, C., Tomlins, S. A. & Chinnaiyan, A. M. Recurrent gene fusions in prostate cancer. *Nature Rev. Cancer* **8**, 497–511 (2008).
2. Clark, J. P. & Cooper, C. S. ETS gene fusions in prostate cancer. *Nature Rev. Urol.* **6**, 429–439 (2009).

3. Bartel, F. O., Higuchi, T. & Spyropoulos, D. D. Mouse models in the study of the Ets family of transcription factors. *Oncogene* **19**, 6443–6454 (2000).
4. Gutierrez-Hartmann, A., Duval, D. L. & Bradford, A. P. ETS transcription factors in endocrine systems. *Trends Endocrinol. Metab.* **18**, 150–158 (2007).
5. Wertz, I. E. et al. Human De-etiolated-1 regulates *c-Jun* by assembling a CUL4A ubiquitin ligase. *Science* **303**, 1371–1374 (2004).
6. Qi, L. et al. TRB3 links the E3 ubiquitin ligase *COP1* to lipid metabolism. *Science* **312**, 1763–1766 (2006).
7. Tomlins, S. A. et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448**, 595–599 (2007).
8. Seo, H. S. et al. LAF1 ubiquitination by *COP1* controls photomorphogenesis and is stimulated by SPA1. *Nature* **423**, 995–999 (2003).
9. Bianchi, E. et al. Characterization of human constitutive photomorphogenesis protein 1, a RING finger ubiquitin ligase that interacts with Jun transcription factors and modulates their transcriptional activity. *J. Biol. Chem.* **278**, 19682–19690 (2003).
10. Yi, C. & Deng, X. W. *COP1* – from plant photomorphogenesis to mammalian tumorigenesis. *Trends Cell Biol.* **15**, 618–625 (2005).
11. Pickart, C. M. Ubiquitin in chains. *Trends Biochem. Sci.* **25**, 544–548 (2000).
12. Baert, J. L. et al. The E3 ubiquitin ligase complex component *COP1* regulates PEA3 group member stability and transcriptional activity. *Oncogene* **29**, 1810–1820 (2010).
13. Hermans, K. G. et al. Truncated *ETV1*, fused to novel tissue-specific genes, and full-length *ETV1* in prostate cancer. *Cancer Res.* **68**, 7541–7549 (2008).
14. Xin, L., Ide, H., Kim, Y., Dubey, P. & Witte, O. N. *In vivo* regeneration of murine prostate from dissociated cell populations of postnatal epithelia and urogenital sinus mesenchyme. *Proc. Natl Acad. Sci. USA* **100** (Suppl. 1), 11896–11903 (2003).
15. Leong, K. G., Wang, B. E., Johnson, L. & Gao, W. Q. Generation of a prostate from a single adult stem cell. *Nature* **456**, 804–808 (2008).
16. Wu, X. et al. Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. *Mech. Dev.* **101**, 61–69 (2001).
17. Shappell, S. B. et al. Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. *Cancer Res.* **64**, 2270–2305 (2004).
18. Ouyang, X. et al. Activator protein-1 transcription factors are associated with progression and recurrence of prostate cancer. *Cancer Res.* **68**, 2132–2144 (2008).
19. Ali, I. U., Schriml, L. M. & Dean, M. Mutational spectra of *PTEN/MMAC1* gene: a tumor suppressor with lipid phosphatase activity. *J. Natl. Cancer Inst.* **91**, 1922–1932 (1999).
20. Yoshimoto, M. et al. Interphase FISH analysis of *PTEN* in histologic sections shows genomic deletions in 68% of primary prostate cancer and 23% of high-grade prostatic intra-epithelial neoplasias. *Cancer Genet. Cytogenet.* **169**, 128–137 (2006).
21. King, J. C. et al. Cooperativity of *TMPRSS2-ERG* with PI3-kinase pathway activation in prostate oncogenesis. *Nature Genet.* **41**, 524–526 (2009).
22. Carver, B. S. et al. Aberrant *ERG* expression cooperates with loss of *PTEN* to promote cancer progression in the prostate. *Nature Genet.* **41**, 619–624 (2009).
23. Cai, C. et al. *ETV1* is a novel androgen receptor-regulated gene that mediates prostate cancer cell invasion. *Mol. Endocrinol.* **21**, 1835–1846 (2007).
24. de Launoit, Y. et al. The Ets transcription factors of the PEA3 group: transcriptional regulators in metastasis. *Biochim. Biophys. Acta* **1766**, 79–87 (2006).
25. Whang, Y. E. et al. Inactivation of the tumor suppressor *PTEN/MMAC1* in advanced human prostate cancer through loss of expression. *Proc. Natl Acad. Sci. USA* **95**, 5246–5250 (1998).
26. Salghetti, S. E., Kim, S. Y. & Tansey, W. P. Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc. *EMBO J.* **18**, 717–726 (1999).
27. Bhatia, K. et al. Point mutations in the *c-Myc* transactivation domain are common in Burkitt's lymphoma and mouse plasmacytomas. *Nature Genet.* **5**, 56–61 (1993).
28. Albert, T., Uhlbauer, B., Kohlhuber, F., Hammersen, B. & Eick, D. Ongoing mutations in the N-terminal domain of *c-Myc* affect transactivation in Burkitt's lymphoma cell lines. *Oncogene* **9**, 759–763 (1994).
29. Clark, H. M. et al. Mutations in the coding region of *c-MYC* in AIDS-associated and other aggressive lymphomas. *Cancer Res.* **54**, 3383–3386 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. White and A. Chinnaiyan for reagents, B. Bolon for pathology support, and D. Dunlap for *in situ* hybridization.

Author Contributions A.C.V. designed and performed *in vitro* experiments. A.C.V., K.G.L., C.Y. and W.-Q.G. designed and performed *in vivo* experiments. K.N. designed and generated the *Cop1* mutant mice. R.V., R.F. and J.-A.H. generated monoclonal antibodies and performed immunohistochemistry. K.O. made constructs. J.L. performed bioinformatics analyses. S.M. and A.P. designed and performed FISH experiments. L.P. and D.A. designed and performed mass spectrometry experiments. S.S.C. and D.M.F. assessed the histopathology of mouse and human tissues. A.C.V., K.G.L., K.N., K.O., J.L., L.P., R.F., S.M., A.P., D.A., D.M.F. and V.M.D. prepared the manuscript and figures. A.C.V., K.G.L., K.N., C.Y., J.L., R.F., S.M., A.P., D.A., I.E.W., W.-Q.G., D.M.F. and V.M.D. contributed to the study design and data analyses.

Author Information Microarray data has been deposited with the Gene Expression Omnibus under accession codes GSE27914. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to V.M.D. (dixit@gene.com).

METHODS

Mice. *Cop1* mutant mice were generated from C2 C57BL/6 embryonic stem (ES) cells electroporated with targeting constructs that: (1) replaced exon 1 coding sequence with a *lacZ* reporter gene or (2) contained exon 3 (encoding amino acids 159–190) flanked by *loxP* sites (Supplementary Fig. 4). Chimeras generated with exon 1 homologous recombinants were bred to C57BL/6-*Gt(ROSA)26Sor^{tm16(Cre)Ar}* mice (TaconicArtemis GmbH) to delete the neomycin selection cassette. Exon 3 homologous recombinants were electroporated with a Cre recombinase expression construct to remove the *loxP*-flanked neomycin selection cassette. All ES clones and mice were confirmed by Southern blotting. *pCAGG.Cre-ER* mice were described previously³⁰. Probasin-Cre4 transgenic mice¹⁶ were backcrossed to C57BL/6N for at least five generations. Mice with conditional *Pten* alleles³¹ were backcrossed to C57BL/6N for at least 10 generations. Pregnant SD rats were from Charles River Laboratories. Athymic nu/nu male mice (6–8 weeks old) were from Harlan Sprague Dawley. The Genentech Institutional Animal Care and Use Committee approved all protocols.

Genotyping. *Cop1* exon 1 primers 5'-GCTACCATTACCAGTTGGTCTGGT GTC-3', 5'-CCAACCCACAAAGTTCAGGGAT-3', and 5'-CTGCATCATGT TGTGTGATTGCAT-3' yield 873 bp wild-type and 532 bp knockout DNA fragments. *Cop1* exon 3 primers 5'-CATTGAAATGATAATTGCAGATTTGGTC-3', 5'-CACCACCCTGCCAGATCTTAAATATAGAT-3', 5'-CAAACCTGTCACA AAATACTATTGTGCTCTC-3' yield 686 bp wild-type, 753 bp floxed, and 452 bp knockout DNA fragments. *Pten* primers 5'-TCCAGAGTTCATACCA GGA-3', 5'-GCAATGGCCAGTACTAGTGAAC-3', 5'-AATCTGTGCATGA AGGGAAC-3' yield ~500 bp wild-type, ~650 bp floxed, and ~300 bp knockout DNA fragments. Cre-specific primers 5'-GCTAAACATGCTTCATCGTCG GTC-3' and 5'-CCAGACCAGGCCAGGTATCTCTG-3' amplified a 582 bp fragment.

Plasmids and cell transfections. Mouse *ETV1* (NP_031986) from a hypothalamus cDNA library was subcloned into pCMVFlag6C (Sigma). Human *ETV1* (NP_004947) was cloned into pCDNA3.1 Myc-HisA (Invitrogen). *ETV4* (NP_001073143), *ETV5* (NP_004445) and *ERG* (NP_001129626) were cloned into pCDNA3 (Invitrogen), which had been modified to contain a C-terminal haemagglutinin (HA) fusion tag. Flag-tagged human *COP1*, *DET1* and *TRAF3* constructs were described previously^{5,32}. Wild-type and RING mutant *COP1* were subcloned into a pHUSH-GW (Invitrogen) inducible vector system to generate stable PC3 clones. HEK293T and BMK cells were transfected with FUGENE 6 (Roche). PC3 and HCC1806 cells were transfected with FUGENE HD (Roche). ON-TARGETplus siRNA oligonucleotides (Dharmacon) were transfected with Lipofectamine2000 and Lipofectamine RNAiMax (Invitrogen) into PC3 and LNCaP cells, respectively. Cell lysates were prepared 48 h after plasmid DNA transfection and 72 h after oligonucleotide transfection.

Affinity purification of complexes containing COP1. Primary baby mouse kidney (BMK) cells were isolated from *Cop1^{fl/Δ1} pCAGG.Cre-ER⁺* mice and immortalized with E1A and dominant negative p53(Δ15–301)³³. Subsequently, cells were grown for 72 h in 100 nM 4-hydroxytamoxifen (Sigma) to generate *Cop1^{Δ3/Δ1}* cells. PCR and western blotting confirmed deletion of the *Cop1* conditional allele. *COP1*-deficient BMK cells were transfected with Flag-COP1, Flag-TRAF3, or empty vector and 24 h later were grown for 4 h in 10 μM MG-132. Cells were lysed in ice-cold buffer (20 mM Tris pH 7.8, 92 mM NaCl, 9 mM MgCl₂, 0.1% Triton X-100, Complete protease inhibitor cocktail (Roche), 10 μM MG-132, phosphatase inhibitor cocktails 1 and 2 (Sigma), 10 U ml⁻¹ DNase I). The soluble lysate was immunoprecipitated with Flag M2 agarose overnight at 4 °C. Agarose beads were washed and bound proteins were eluted with 0.5 mg ml⁻¹ 3×-Flag peptide (Sigma).

Mass spectrometry. *COP1*-binding proteins concentrated by ultrafiltration (YM3, Microcon) were reduced and alkylated for SDS-PAGE. Resolved proteins were stained with Coomassie blue R-250 and subjected to in-gel trypsin digestion³⁴. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis of the tryptic digests was performed on a hybrid linear ion trap Fourier transform ion cyclotron resonance mass spectrometer (LTQ-FT; Thermo Fisher) coupled to a nano-flow HPLC system (MDLC; Eksigent) in a vented column configuration³⁵. Mass spectral data were acquired using a data-dependent method comprised of one full MS scan (400–2,000 *m/z*) followed by product ion scans on the five most abundant ions detected. Mascot software (Matrix Science) was used to search the Swiss-Prot database. Results were displayed with Scaffold (Proteome Software), protein and peptide probability filters both set to 95%.

Bioinformatics. The consensus *COP1* binding motif [D/E](x)xxVP[D/E] was extracted from an alignment of *Arabidopsis thaliana* HY5, HYH, STO, STH; *Mus musculus* TRB3 and CRT2; *Homo sapiens* c-JUN, JUNB, and JUND^{5,6,9,36–38}.

mRNA expression profiling was carried out on Affymetrix GeneChip Human Genome U133 Plus 2.0 Array following the manufacturer's protocol (GEO accession GSE27914). Expression summary values for all probe-sets were calculated

using the RMA algorithm as implemented in the affy package from Bioconductor. Probe sets 234950_s_at and 221911_at were chosen to assess *COP1* and *ETV1*, respectively (Fig. 1a). Breast cancer cell lines HCC1806 and HCC1937 lacked expression of *DET1* mRNA and had very low *ETV1* mRNA (data not shown). Statistical analyses of differentially expressed genes were performed using linear models and empirical Bayes moderated statistics as implemented in the limma package from Bioconductor.

DNA copy number data for human *COP1* in prostate cancer samples were extracted from two public Affymetrix SNP array data sets (GEO accession GSE12702, *n* = 20; GSE19399, *n* = 87, including data downloaded from <http://www.broadinstitute.org/tumorscape/pages/portalHome.jsf>), and one Agilent Human Genome CGH 244A dataset generated by Genentech (GSE20393, *n* = 59). All raw data were processed with the Genentech internal data analysis pipeline. For the Affymetrix SNP array data, array intensity signal CEL files were first processed by dChip using the PM/MM difference model and invariant set normalization, and normalized with data for normal samples from the Affymetrix website (<http://www.affymetrix.com>). Agilent CGH array data were first processed by Feature Extraction Software from Agilent. All processed copy numbers were then centred to a median of 2 and segmented. Copy number values for specific genes were calculated as the mean copy number value for the probe sets bounding the gene location and all intervening probe sets using the segmented data.

siRNAs. *COP1* siRNA1 5'-CUACAAGGAUGUCUCGUAU-3'; *COP1* siRNA2 5'-GCUAAUGUGUGUCUGUGUAU-3'; *COP1* siRNA3 5'-GAAUUGGUAUGA AGGUUA-3'; *COP1* siRNA4 5'-CAUAAAGAACCGUUAUGUA-3'; *ETV1* siRNA1 5'-GAACAGCCCUUUAUUUA-3'; *ETV1* siRNA2 5'-CAACGAAG GCUACGUGUAU-3'; *ETV1* siRNA3 5'-UCUCCAAACUCAACUCAU-3'; *ETV1* siRNA4 5'-GAGAAAUUGUAACGAGAAA-3'; *ETV4* siRNA1 5'-GGGC AGAGCAACGGAAUUU-3'; *ETV4* siRNA2 5'-GAAUGGAGUUAACGCUC AU-3'; *ETV4* siRNA3 5'-GGACUUCGCCUACGACUCA-3'; *ETV4* siRNA4 5'-GAUGAAAGCCGGAUACUUG-3'; *ETV5* siRNA1 5'-CCGAAAGGCUUUGC UUACUA-3'; *ETV5* siRNA2 5'-CGGCAAAUGUCAGAACCUA-3'; *ETV5* siRNA3 5'-GAGAUAAUCGCCCCAGUUA-3'; *ETV5* siRNA4 5'-GGAAUUCU CGAUCUGAGGA-3'; *DET1* siRNA1 5'-GUAGUAAACACUGCGAGUCA-3'; *DET1* siRNA3 5'-CAAGUACACUAGUGAGGAU-3'; *c-JUN* siRNA6 5'-GAA CAGGUGGCACAGCUUA-3'; *c-JUN* siRNA7 5'-GAAACGACCUUUAUGA CG-3'; Non-targeting siRNAs were from Dharmacon (catalogue D-001810-01-20; D-001810-02-20; D-001810-10-05; D-001810-01-05).

Immunoprecipitations and western blotting. Cells were lysed in 20 mM HEPES pH 7.2, 2 mM EGTA, 5 mM EDTA, 30 mM NaF, 60 mM β-glycerophosphate, 20 mM sodium pyrophosphate, 1 mM Na₃VO₄, 1% Triton X-100. Soluble lysate was immunoprecipitated with Flag M2 agarose or with the indicated antibodies coupled to Protein A/G agarose beads (Pierce Biotechnology). Beads were washed extensively in lysis buffer containing 0.5 M NaCl and then once more in straight lysis buffer before elution in LDS-sample buffer (Invitrogen) containing 1% 2-mercaptoethanol. To detect ubiquitinated MYC-ETV1, soluble lysate was supplemented with 1% SDS and heated at 95 °C for 10 min. Denatured lysate was diluted 20-fold in lysis buffer and then immunoprecipitated with 9E10 MYC agarose (Clontech).

Antibodies used for immunoblotting recognized β-actin (Novus Biologicals, NB600-501), c-JUN (Epitomics, 1254-1), *COP1* (Genentech, 28A4), *DET1* (Genentech, 3G5), *ETV1* (Abcam, Ab36788), *ETV1* (Genentech rat monoclonal 13G11), *ETV4* (Genentech rabbit polyclonal Y771, which recognizes amino acids 2–199 of mouse *ETV4*), *ETV4* (Genentech rat monoclonal 5F8), Flag (Sigma, M2), GAPDH (Cell Signalling, 14C10), GFP (Invitrogen, A11122), HA (Sigma HA-7), K48- or K63-linked polyubiquitin (Genentech Apu2.07 and Apu3.A8)³⁹, and MYC (Genetex, GTX21261). Unless indicated, *ETV1* was blotted with Ab36788 throughout.

Antibodies used for immunoprecipitation detected *COP1* (28A4) and *ETV1* (Genentech rabbit polyclonal antibodies Y713 and Y714, which recognize amino acids 268–477 of mouse *ETV1*). Quantification of western blot signals was performed on a Typhoon scanner (GE Healthcare) following chemiluminescence detection with ECLplus (GE Healthcare).

Cell invasion. PC3 clonal cell lines stably expressing doxycycline-inducible *COP1* variants were grown overnight in medium lacking fetal bovine serum (FBS) and in the presence of 0.03 μg ml⁻¹ of doxycycline or PBS vehicle control. In the morning, cells were trypsinized, resuspended in the same medium, counted, and then assayed for their invasiveness. Cells (10⁵) were seeded in quadruplicate in 24-well transwell plates (8 μm pores, Fluoroblok PET membranes (BD Falcon)) that were previously coated with rat tail collagen type 1. Medium containing 10% FBS was added to lower compartment. After 23–24 h live cells were stained with calcein (1 μg ml⁻¹ for 1 h), imaged and quantified on an ImageXpress microscope device (Molecular Devices). In Fig. 2h–j, cells were cotransfected with GFP expressing

vector and the desired expression vectors 24 h before serum deprivation. GFP-expressing cells that migrated into the lower chamber were imaged.

Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR).

Total cellular RNA was prepared with a RNeasy Plus kit (Qiagen) and subject to DNase I digestion. qRT-PCR reactions were performed in 384-well plates, with 10 to 40 ng of RNA per reaction on a 7900HT Fast Real-Time PCR System (Applied Biosystems). Taqman gene expression assays (Applied Biosystems): *COP1*, Hs00375437_m1; *DET1*, Hs00894490_m1; *ETV1*, Hs00231877_m1; *ETV4*, Hs00385910_m1; *ETV5*, Hs00231790_m1; *MMP1*, Hs00899658_m1; *MMP7*, Hs01042796_m1, *RPLP0*, 4326314E. Control reactions lacked reverse transcriptase and had C_t values at least 3 units higher than reactions performed with reverse transcriptase.

MMP1 and MMP7 protein quantification. Forty eight hours following transfection, LNCaP cells were grown overnight in serum-free medium. Total MMP1 and MMP7 proteins in conditioned medium were quantified by ELISA with SensoLytic MMP1 (AnaSpec) and Quantikine MMP7 (R&D Systems).

Prostate regeneration model. Rat UGM stromal cell isolation¹⁵ and the prostate regeneration assay were described previously^{14,40}. Prostate from 8 to 10 weeks old mice were dissociated and live cells enriched with a Dead Cell Removal Kit (Miltenyi Biotec). Prostate cells (100,000 cells per graft) were mixed with UGM stromal cells (250,000 cells per graft) in 3 mg ml⁻¹ collagen type I (20 µl per graft), incubated at 37 °C for 1 h to allow collagen gelation, and overlaid with prostate culture medium (DMEM supplemented with 10% FBS, 2 mM glutamine, 10 µg ml⁻¹ insulin, 5.5 µg ml⁻¹ transferrin, 6.7 ng ml⁻¹ selenium, 1 nM testosterone (Innovative Research of America), 100 U ml⁻¹ penicillin and 100 mg ml⁻¹ streptomycin). Gels were incubated overnight at 37 °C with 10⁹ plaque forming units per ml of Ad5-CMV-Cre-GFP virus (Baylor College of Medicine), washed in culture medium, and grafted under the renal capsule of 6–8 weeks old athymic nu/nu mice together with a subcutaneous 90-day slow-release testosterone pellet (12.5 mg per pellet per mouse; Innovative Research of America). Grafts were harvested 12 weeks after implantation.

Immunohistochemistry. Immunohistochemical analyses of regenerated prostate structures were performed as described¹⁵ with antibodies to CK18 (Abcam, C-04), Ki67 (BD Biosciences, clone B56), CK14 (Covance, AF64), β 1-integrin (Chemicon, clone MB1.2), and probasin (Santa Cruz Biotechnology, M-18). Percentages of positive cells were calculated by assessing at least 1,000 cells per genotype stained for CK18, 900 cells for Ki67, and 300 cells for CK14.

Formalin-fixed paraffin-embedded mouse prostates were cut at 4 µm, pre-treated with Target Antigen Retrieval buffer (DAKO) followed by KPL peroxidase blocking solution (Kirkegaard and Perry Laboratories) and avidin/biotin blocking kit (Vector Labs). TNB blocking buffer (Perkin Elmer) or 10% goat serum/3% BSA/PBS was used before incubation with either 1 µg ml⁻¹ rat ETV1 antibody (Genentech, Clone 1H2), 0.125 µg ml⁻¹ hamster COP1 antibody (Genentech, Clone 1D10), 0.3 µg ml⁻¹ rabbit JUN antibody (Epitomics, Clone E254) or at 1:200 with rabbit Ki67 antibody (Thermo Scientific, Clone SP6) for 60 min at room temperature. Species-appropriate biotinylated secondary antibody (Vector Labs), followed by a streptavidin-HRP reagent from TSA kit (for ETV1 and COP1 IHC; Perkin Elmer) or ABC Elite HRP Reagents (for c-JUN and Ki67 IHC; Vector Labs) was applied. Biotinylated TSA amplification reagents (Perkin Elmer) were used to visualize ETV1 and COP1 immunostaining. Sections were treated with metal enhanced DAB colorimetric peroxidase substrate (Thermo Scientific), then Myer's haematoxylin (Rowley Biochemical Institute) counterstain.

Rabbit ETV4 antibody (Lifespan Biosciences) staining of mouse prostate was performed on the Ventana Discovery XT Platform at 5 µg ml⁻¹ with no antigen retrieval using the anti-murine-OMNIMAP-HRP Kit and Ventana DAB colorimetric reagents. Sections were counterstained with Ventana Hematoxylin II reagent. Experiments to validate the 1H2 rat anti-ETV1, Ab36788 rabbit anti-ETV1 and the Lifespan anti-ETV4 antibodies are presented (Supplementary Fig. 1e, f and Supplementary Fig. 10).

In situ hybridization (ISH). Radioactive *in situ* hybridization was performed as previously described⁴¹. Probes were prepared by PCR with the following primers: human *COP1* (5'-GGGCTCATCAACTCCTACGA-3'; 5'-GAGAACTGCCAC TGAAACCTG-3'); human *ETV1* (5'-GAATCTTTGTTTTATTCTGTGT-3'; 5'-CAGAGTCCAAAATTGTGCCCTC-3'). The slides were exposed for 5 weeks, developed, and counterstained with haematoxylin and eosin.

Fluorescence in situ hybridization (FISH). A bacterial artificial chromosome (BAC) contig comprising three overlapping clones, RP11-102E20, CTD-3127J24, and RP11-415M4 was used as a *COP1* probe. A *CEP1* probe (Vysis/Abbott Laboratories) was also used. FISH probes for identifying ERG and ETV1 translocations were provided by A. M. Chinnaiyan and used as described⁴². BAC clone DNA was extracted by standard methods⁴³ and directly labelled with Spectrum Orange by nick translation (Vysis/Abbott Laboratories). FISH to normal human metaphases (Abbott Laboratories) confirmed the genomic location of the BAC clones. FISH on cytogenetic preparations and formalin-fixed paraffin-embedded tissue was performed as described⁴⁴. *COP1* copy number was evaluated by counting spots in a range from 50 to 100 non-overlapped, intact interphase nuclei per tumour tissue core. 4',6-diamidino-2-phenylindole, dihydrochloride staining of nuclei with reference to the corresponding haematoxylin- and eosin-stained tissue identified the areas of adenocarcinoma. Based on hybridization in control normal cells (data not shown), hemizygous deletion of *COP1* was defined as >30% (mean + 3 s.d. in non-neoplastic cells) of tumour nuclei containing one *COP1* locus signal and by the presence of CEP1 signals.

30. Hayashi, S. & McMahon, A. P. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. *Dev. Biol.* **244**, 305–318 (2002).
31. Wang, S. *et al.* Prostate-specific deletion of the murine Pten tumor suppressor gene leads to metastatic prostate cancer. *Cancer Cell* **4**, 209–221 (2003).
32. Kayagaki, N. *et al.* DUBA: a deubiquitinase that regulates type I interferon production. *Science* **318**, 1628–1632 (2007).
33. Shaulian, E., Zauberman, A., Ginsberg, D. & Oren, M. Identification of a minimal transforming domain of p53: negative dominance through abrogation of sequence-specific DNA binding. *Mol. Cell. Biol.* **12**, 5581–5592 (1992).
34. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols* **1**, 2856–2860 (2007).
35. Yi, E. C., Lee, H., Aebersold, R. & Goodlett, D. R. A microcapillary trap cartridge-microcapillary high-performance liquid chromatography electrospray ionization emitter device capable of peptide tandem mass spectrometry at the attomole level on an ion trap mass spectrometer with automated routine operation. *Rapid Commun. Mass Spectrom.* **17**, 2093–2098 (2003).
36. Holm, M., Hardtke, C. S., Gaudet, R. & Deng, X. W. Identification of a structural motif that confers specific interaction with the WD40 repeat domain of *Arabidopsis* COP1. *EMBO J.* **20**, 118–127 (2001).
37. Ang, L. H. *et al.* Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of *Arabidopsis* development. *Mol. Cell* **1**, 213–222 (1998).
38. Dentin, R. *et al.* Insulin modulates gluconeogenesis by inhibition of the coactivator TORC2. *Nature* **449**, 366–369 (2007).
39. Newton, K. *et al.* Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell* **134**, 668–678 (2008).
40. Cunha, G. R. & Lung, B. The possible influence of temporal factors in androgenic responsiveness of urogenital tissue recombinants from wild-type and androgen-insensitive (Tfm) mice. *J. Exp. Zool.* **205**, 181–193 (1978).
41. Jubb, A. M., Pham, T. Q., Frantz, G. D., Peale, F. V. Jr & Hillan, K. J. Quantitative *in situ* hybridization of tissue microarrays. *Methods Mol. Biol.* **326**, 255–264 (2006).
42. Mehra, R. *et al.* Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. *Mod. Pathol.* **20**, 538–544 (2007).
43. O'Brien, C. *et al.* Functional genomics identifies ABCC3 as a mediator of taxane resistance in HER2-amplified breast cancer. *Cancer Res.* **68**, 5380–5389 (2008).
44. Pandita, A., Aldape, K. D., Zadeh, G., Guha, A. & James, C. D. Contrasting *in vivo* and *in vitro* fates of glioblastoma cell subpopulations with amplified EGFR. *Genes Chromosom. Cancer* **39**, 29–36 (2004).

Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA

Dong Wang^{1*}, Ivan Garcia-Bassets^{2,3*}, Chris Benner^{1*}, Wenbo Li², Xue Su^{2,4}, Yiming Zhou⁵, Jinsong Qiu¹, Wen Liu^{2,4}, Minna U. Kaikkonen¹, Kenneth A. Ohgi², Christopher K. Glass¹, Michael G. Rosenfeld² & Xiang-Dong Fu¹

Mammalian genomes are populated with thousands of transcriptional enhancers that orchestrate cell-type-specific gene expression programs^{1–4}, but how those enhancers are exploited to institute alternative, signal-dependent transcriptional responses remains poorly understood. Here we present evidence that cell-lineage-specific factors, such as FoxA1, can simultaneously facilitate and restrict key regulated transcription factors, exemplified by the androgen receptor (AR), to act on structurally and functionally distinct classes of enhancer. Consequently, FoxA1 downregulation, an unfavourable prognostic sign in certain advanced prostate tumours, triggers dramatic reprogramming of the hormonal response by causing a massive switch in AR binding to a distinct cohort of pre-established enhancers. These enhancers are functional, as evidenced by the production of enhancer-templated non-coding RNA (eRNA⁵) based on global nuclear run-on sequencing (GRO-seq) analysis⁶, with a unique class apparently requiring no nucleosome remodelling to induce specific enhancer–promoter looping and gene activation. GRO-seq data also suggest that liganded AR induces both transcription initiation and elongation. Together, these findings reveal a large repository of active enhancers that can be dynamically tuned to elicit alternative gene expression programs, which may underlie many sequential gene expression events in development, cell differentiation and disease progression.

The wide diversity of mammalian cells is determined by a large repertoire of constitutive and inducible genes, which are regulated by general and cell-type-specific transcription factors and cofactors through regulatory genomic elements^{7,8}. Recent studies reveal that gene promoters are marked by tri-methylated H3K4 (H3K4me3) and distal regulatory elements are often associated with mono-methylated H3K4 (H3K4me1)^{1,2}. Because these H3K4me1-positive, H3K4me3-negative regions exhibit striking cell-type specificity^{1,2}, we used this signature to characterize potential enhancers in prostatic LNCaP cells in which one of key regulatory transcriptional programs is mediated by the AR. We identified by chromatin immunoprecipitation (ChIP)-sequencing 14,283 H3K4me3-marked and 51,544 H3K4me1-marked loci in androgen (5 α -dihydrotestosterone, (DHT))-treated LNCaP cells, among which 43,565 loci are uniquely marked by H3K4me1, largely localized distal to annotated transcriptional start sites (TSSs) (94%), and associated with other marks linked to enhancer activities (Fig. 1a).

De novo DNA motif analysis revealed several highly enriched motifs, particularly the forkhead motif (Fig. 1b). Using a specific antibody against FoxA1, a major FOX family member expressed in LNCaP cells and normal prostate gland^{9–11} (Supplementary Fig. 1), we identified 33,426 FoxA1-bound sites, which extensively overlap with distal H3K4me1-marked regions (Fig. 1c and Supplementary Fig. 2a; see on *KLK3* enhancer¹² in Supplementary Fig. 2b). RNA profiling supports the functional relevance of these FoxA1/H3K4me1 loci, as genes responsive to *FOXA1* short interfering RNA (siRNA) are

located more proximally to FoxA1/H3K4me1-marked loci than non-responsive genes (Fig. 1d and Supplementary Fig. 3).

FoxA1 has been characterized as a ‘pioneer’ factor to facilitate DNA binding by other sequence-specific transcription factors^{9,13–16} and ‘translate’ H3K4me1/me2 into AR-mediated gene expression⁹. Comparing the profile of H3K4me1 and H3K27ac before and after *FOXA1* knock-down, we detected three classes of FoxA1-binding sites based on the H3K4me1 signal exhibiting reduced (~22%), relatively unaffected (~74%) or even increased (~3.4%) levels over candidate enhancers (Fig. 1e–g and Supplementary Fig. 4). RNA profiling analysis agrees with the functional significance of these selective FoxA1 effects, revealing more downregulated genes in the first class, roughly equal numbers of up- or downregulated genes in the second and more upregulated genes in the third (Fig. 1h), suggesting a contribution of FoxA1 to ‘writing’ and ‘reading’ the ‘histone code’ on different enhancer cohorts, in line with its critical function in prostate gland development^{10,11}.

The rationale for our experimental strategy to use RNA interference (RNAi) to study FoxA1-regulated enhancer network is the association of decreased *FOXA1* expression with castration-resistant, poor prognostic prostate tumours (Supplementary Fig. 5). In LNCaP cells, *FOXA1* RNAi enhanced cell entrance to S phase with reduced hormone (Fig. 2a). To understand the mechanistic basis for elevated hormone responsiveness, we mapped AR-binding sites, identifying 3,115 high confident loci with approximately 65% co-incident with H3K4me1. *De novo* motif analysis revealed highly enriched elements for both AR and FoxA1, including a composite motif consisting of a FOX motif and AR regulatory element (ARE) half site, suggesting ternary complex formation on these sites (Fig. 2b). Indeed, 1,684 AR-bound loci (54% of total) are co-occupied by FoxA1 in DHT-treated LNCaP cells and FoxA1 appears to bind to most of these sites (~70%) before hormone treatment (Supplementary Fig. 6).

The conundrum is that, although FoxA1 is known to facilitate AR binding on several DHT-responsive genes⁹, *FOXA1* RNAi actually markedly elevated, rather than diminished, the DHT response (Fig. 2a). We found that approximately 60% of the original AR binding events were ‘expectedly’ lost in response to *FOXA1* RNAi, which we refer to as the ‘lost’ AR program (Fig. 2c, d). We refer to the remaining approximate 40% of AR binding events as the ‘conserved’ AR program, which often exhibited enhanced AR binding. Strikingly, we detected a massive gain of 10,869 new AR binding loci, referred to as the ‘gained’ AR program (Fig. 2c, d). We extensively validated each of these AR programs by conventional ChIP–quantitative PCR (qPCR) (Fig. 2e). This induced AR reprogramming appears to be qualitatively and quantitatively distinct from reported AR re-targeting on androgen-resistant LNCaP-abl cells compared with parental LNCaP cells¹⁷ and is in sharp contrast to FoxA1-dependent genomic targeting of the oestrogen receptor- α (ER- α) in breast cancer MCF7 cells¹⁸. In concert with such massive AR reprogramming, we observed corresponding

¹Department of Cellular and Molecular Medicine, School of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0651, USA. ²Department of Medicine, Howard Hughes Medical Institute, School of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0651, USA. ³Department of Medicine, Division of Endocrinology and Metabolism, School of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0651, USA. ⁴Graduate Program in Biology, School of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0651, USA. ⁵Digomics LLC, PO Box 30, Malden, Massachusetts 02148, USA.

*These authors contributed equally to this work.

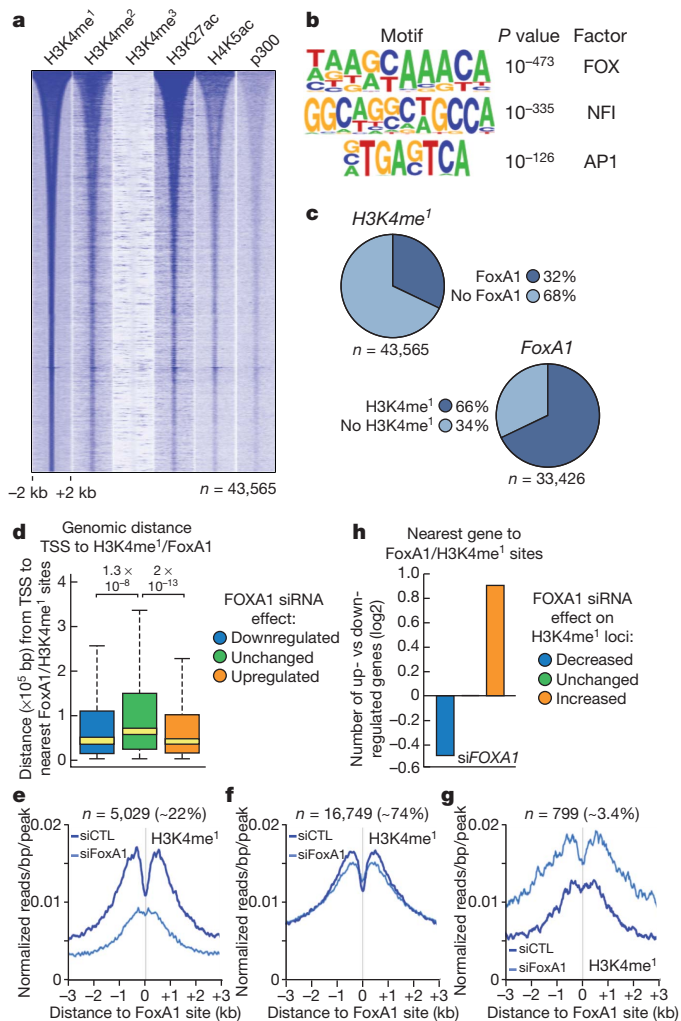


Figure 1 | FoxA1 contributes to the enhancer code in prostate cancer cells.

a, Distribution of histone marks within ± 2 -kb windows around distinct genomic regions ($n = 43,565$) marked by H3K4me1, but not H3K4me3, in androgen (DHT)-stimulated LNCaP cells. The ChIP-seq data sets for H3K4me1, H3K4me2, H3K4me3, H3K27ac, H4K5ac and p300 were each aligned with respect to the centre of the H3K4me1 signal and sorted by the length of H3K4me1-marked regions. **b**, Top-enriched DNA motifs with significant P values and prospective families of DNA binding transcription factors identified by *de novo* motif analysis of non-promoter regions marked by H3K4me1. **c**, Percentage of H3K4me1-marked regions that show FoxA1 binding events (top panel) and percentage of FoxA1-binding sites that are marked by H3K4me1 (bottom panel). Note that H3K4me1-marked regions tend to be broad, but FoxA1-binding sites are discrete; as a result, many H3K4me1-positive regions may contain more than one FoxA1-binding site. **d**, Genomic distance from FoxA1/H3K4me1-positive loci to the nearest TSS of genes in response to FOXA1 knockdown. Outliers were omitted from box plots. P values indicate the significance in pair-wise comparisons. **e–g**, Three classes of FoxA1/H3K4me1-positive loci according to the response in levels of H3K4me1 to FOXA1 knockdown: greater than 1.5-fold decrease (**e**), no significant change (**f**) and greater than 1.5-fold increase (**g**). **h**, Ratio (\log_2) of up- and downregulated genes in each H3K4me1-responsive category in **e–g**. CTL, control.

changes in gene expression in each of three AR programs (Fig. 2f, g and Supplementary Fig. 7). The newly induced AR expression program is also linked to AR binding events (Fig. 2h), suggesting a direct gain-of-function on DHT-responsive genes, as illustrated on SOX9 and other genes (Supplementary Fig. 8), which have been previously documented to play critical roles in cancer progression^{19,20}. Because we also observed an approximate threefold elevation of AR expression in FOXA1 RNAi-treated cells (Supplementary Fig. 9a), we tested the possibility that increased AR expression might trigger these effects. We found that AR overexpression alone was insufficient to induce AR reprogramming (Supplementary Fig. 9b).

To explore the mechanism for AR reprogramming, we determined FoxA1 binding on different AR programs. We found that the gained AR program is largely devoid of FoxA1, whereas FoxA1 is present in more than half of the lost and conserved AR programs (Supplementary Fig. 10). This raises the possibility that FoxA1 may facilitate AR binding to its original binding program, but trans-repress AR from binding to other genomic regions that lack FoxA1-binding sites in the gained program, a strategy frequently used by other transcription activators²¹.

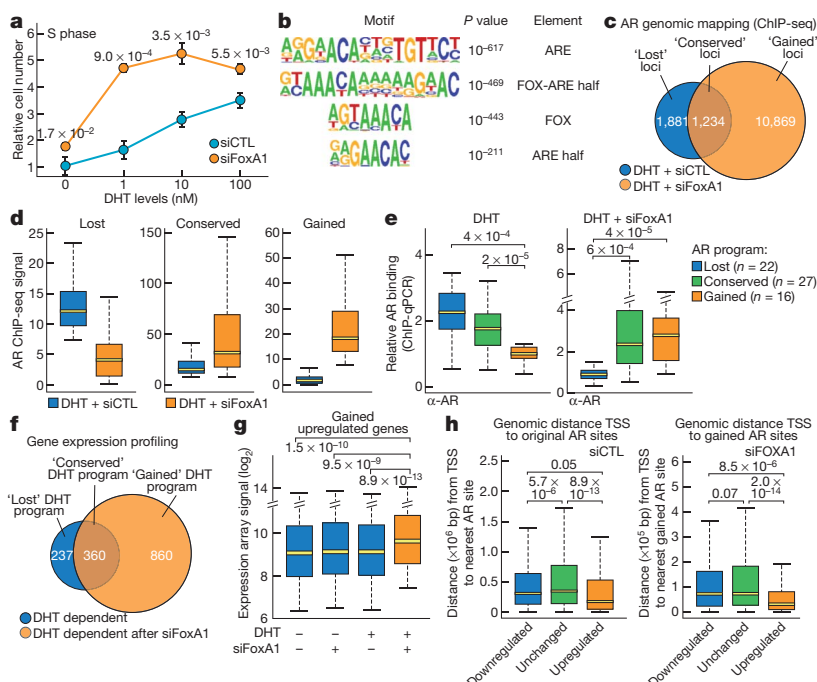


Figure 2 | AR reprogramming and induced alternative hormonal response.

a, FOXA1 siRNA-induced cell progression to S phase. Relative numbers of propidium-iodide-labelled cells in S-phase at different DHT concentrations were determined by FACS. The P value for the difference detected at each hormonal level is indicated; mean \pm s.e.m. is based on four independent experiments. **b**, Top-enriched motifs associated with AR-occupied loci ($n = 3,115$). **c**, Comparison between genome-wide AR binding programs before and after FOXA1 knockdown in DHT-treated LNCaP cells. **d**, Quantitative levels of AR binding in the 'lost', 'conserved' and 'gained' programs. Outliers were omitted from box plots. **e**, ChIP-qPCR-validated AR binding events on randomly selected loci from the lost ($n = 22$), conserved ($n = 27$) and gained ($n = 16$) programs. **f**, Microarray analysis of DHT-induced genes before and after FOXA1 knockdown. **g**, Quantitative analysis of gained androgen upregulated genes based on microarray analysis in **f**. Outliers were omitted from box plots. **h**, Genomic distance of androgen-responsive genes from TSS to the nearest AR-binding site in the original and gained AR-binding programs.

Indeed, as previously reported²², FoxA1 overexpression squelched ARE-driven transcription in transfected HEK293 cells (Supplementary Fig. 11), which is consistent with the ability of AR to interact with FoxA1 directly²³. This mechanism appears to be exploited during tumour progression because an AR mutation identified in advanced prostate tumours lacks part of the hinge domain important for interactions with FoxA1, its ability to interact with FoxA1, and became resistant to FoxA1-mediated trans-repression (Supplementary Fig. 11b, c). Furthermore, our functional analysis indicates that the missing AR ligand-binding domain also contributes to AR:FoxA1 interactions (Supplementary Fig. 12). Interestingly, similar AR truncations have also been reported to result from alternative splicing, gene rearrangement and/or calpain-mediated cleavage (Supplementary Fig. 13). Based on these findings, we propose that FoxA1 regulates AR genomic targeting by simultaneously anchoring AR to cognate loci and restricting AR from other ARE-containing loci in the human genome.

To understand how reprogrammed AR binding is translated to altered hormonal response, we took advantage of the recently established GRO-seq⁶ to detect the functional relationship between AR binding and hormone-induced gene expression. This powerful genome-wide interrogation of ongoing transcription detected a broad scope of nascent RNAs. We uncovered 28,318 transcripts with 15,656 annotated and 12,662 unannotated transcripts, among which 450 coding and 347 unannotated transcripts were induced more than 1.5-fold

with even just 1 h DHT treatment (Supplementary Fig. 14). The TSSs of GRO-seq defined transcripts are typically marked by H3K4me3 and H3K27ac (Supplementary Fig. 15a, b). Importantly, GRO-seq also detected non-coding RNAs from a subset of H3K4me1-positive, H3K4me3-negative regions (Supplementary Fig. 15c). As illustrated on the enhancer of the *KLK3* transcription unit (Fig. 3a), these eRNAs are largely symmetrical and bidirectional (see additional examples on other well-known hormone regulated genes, such as *PMEPA1* and *KLK2* in Supplementary Fig. 16). Interestingly, we often detected a large amount of nascent RNA before DHT treatment, particularly near their TSSs (for example, *KLK3*); DHT not only enhanced the expression of these nascent RNAs, but also allowed the extension of transcription towards the end of the gene (Fig. 3a and Supplementary Fig. 16). We estimated that approximately 79% of the transcription units induced by liganded AR are regulated at the level of transcriptional initiation, whereas approximately 21% appear to be primarily regulated at the level of elongation (Supplementary Fig. 17).

The ability to detect regulated eRNA expression allowed us to analyse different AR programs during transcriptional reprogramming. In the presence of FoxA1, DHT enhanced eRNA expression from AR-bound enhancers in both the lost and conserved AR programs. In contrast, a basal level of eRNAs was detectable on the gained program, but was independent of the hormone treatment, indicating that these are pre-established enhancers (Fig. 3b). In response to *FOXA1* RNAi,

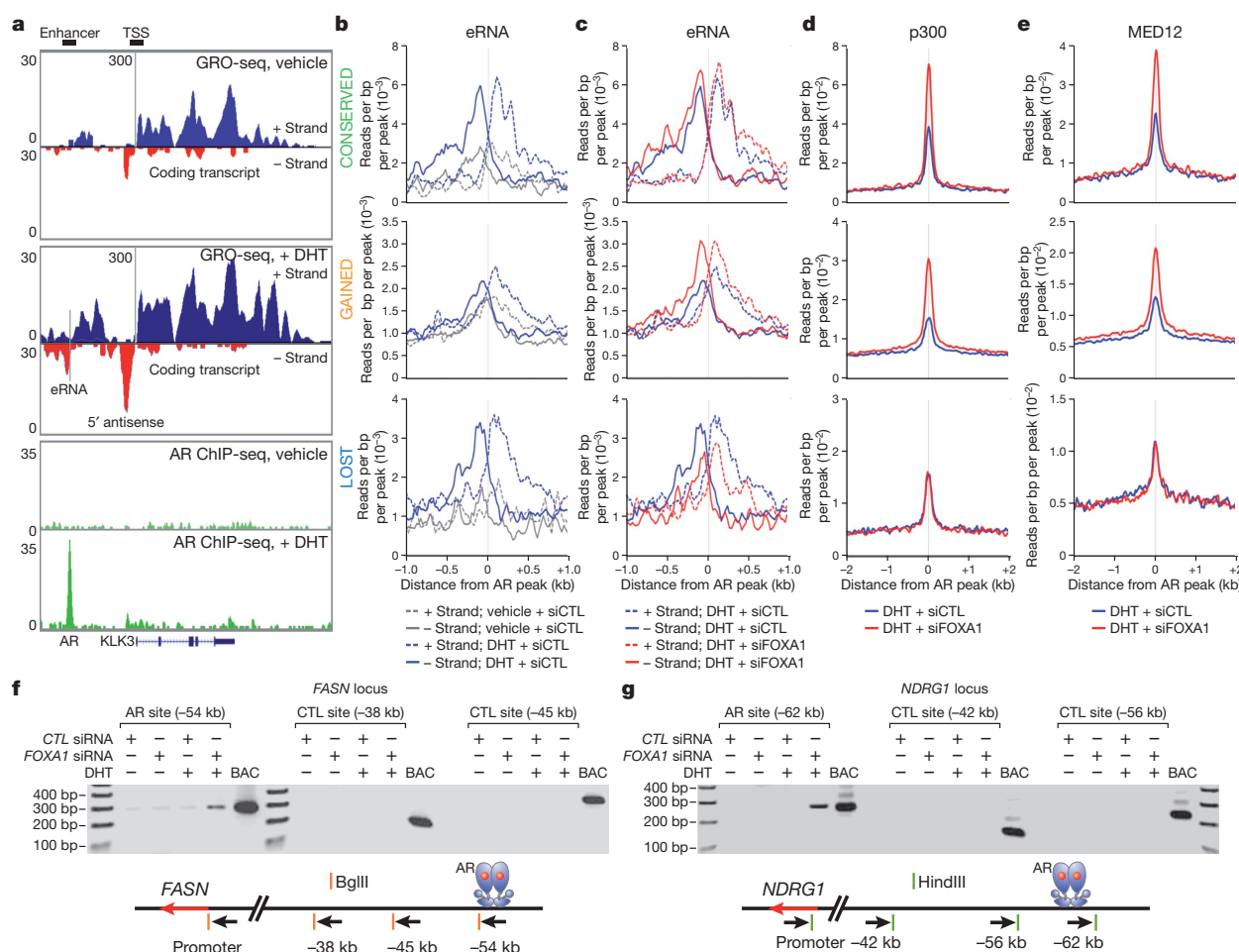


Figure 3 | Transcriptional response on individual enhancer programs to *FOXA1* downregulation. **a**, Display of nascent RNA detected by GRO-seq on the *KLK3* locus. The DHT-induced AR binding is shown at bottom as a reference. **b**, **c**, Induction of eRNA by DHT (**b**) or *FOXA1* knockdown in DHT-treated LNCaP cells (**c**). The eRNA levels under different conditions (indicated at bottom) are separately displayed on three AR-binding programs. **d**, **e**, Effects

of FoxA1 on binding of p300 (**d**) and Med12 (**e**) in each AR program in DHT-treated LNCaP cells. **f**, **g**, Long-distance interaction between gene promoter and AR-bound site was determined by the 3C assay on two representative gene loci selected from the conserved and gained AR programs. Negative controls at shorter distances and a positive control with the corresponding BAC in the region are included in each case.

the expression of eRNAs was diminished from the lost program, but modestly or dramatically enhanced from the conserved and gained programs, respectively (Fig. 3c). The DHT-induced nascent transcripts (detected by GRO-seq) and steady-state RNAs (detected by microarrays) best predict direct target genes by liganded AR, as they show the shortest distance (<50 kilobases (kb)) to nearby AR-binding sites compared with genes identified by either criterion alone (Supplementary Fig. 18), indicating that AR-activated enhancers marked by increased eRNA are responsible for activation of nearby coding transcription units.

In concert with differential eRNA expression, we also observed corresponding changes in levels of another mark in the final step of enhancer activation⁴, specifically p300, on both conserved and gained AR programs (Fig. 3d). Interestingly, enhancers in the lost AR program continued to exhibit significant p300 binding, even after AR binding and eRNA expression were diminished in *FOXA1* knockdown cells (Fig. 3c, d). The transcription mediator Med12 has recently been suggested to mediate enhancer–promoter looping²⁴. We tested Med12 binding on individual AR programs, finding that it exhibited an identical binding pattern to p300 (Fig. 3e). Enhanced Med12 binding on the conserved and gained programs after *FOXA1* knockdown suggests elevated or newly activated enhancer–promoter interactions. This was demonstrated by the 3C assay on two representative genes where *FOXA1* knockdown either enhanced (on the *FASN* locus from the conserved AR program) or create new (on the *NDRG1* locus in the gained AR program) long-range interactions between AR-bound enhancers and specific gene promoters in DHT-treated cells (Fig. 3f,

g and Supplementary Fig. 19). These data strongly suggest that the induction of eRNAs, rather than binding of either p300 or Med12, is the most precise mark of the final, functional looping between an activated enhancer and its regulated gene promoter.

Addressing the structural basis for different functional classes of AR enhancers, we note that the distinct profiles of H3K4me1 and H3K27ac on the lost, conserved and gained AR programs and *FOXA1* RNAi had little effect on these profiles (Fig. 4a, b and Supplementary Fig. 20). The histone marks H3K4me1 and H3K27ac around the lost and conserved AR programs exhibit a bimodal distribution, which is particularly pronounced on the lost program (Fig. 4a, bottom panel). The DNA-binding sites in the lost AR program are actually significantly less enriched in canonical AREs, which may render AR binding on these sites particularly dependent on FoxA1, whereas both the conserved and gained AR programs are associated with nearly perfect palindromic, canonical AREs (Supplementary Fig. 21), explaining why AR is able to target those sites in a FoxA1-independent manner. Strikingly, the gained AR-binding sites are coincident with sharp H3K4me1 and H3K27ac peaks (Fig. 4a, b, middle panels), suggesting a distinct nucleosome architecture underlying the gained AR program.

A recent study has suggested that AR binding leads to dynamic dismissal of a central, H2A.Z-containing nucleosome, being replaced by two flanking H3K4me2-marked nucleosomes²⁵. We found that the lost AR program was largely devoid of a ‘central’ nucleosome even before AR binding (Fig. 4c, bottom panel). The conserved AR program exhibited DHT-induced switch from the central H3K4me2-marked nucleosome to two flanking H3K4me2-marked nucleosomes, which

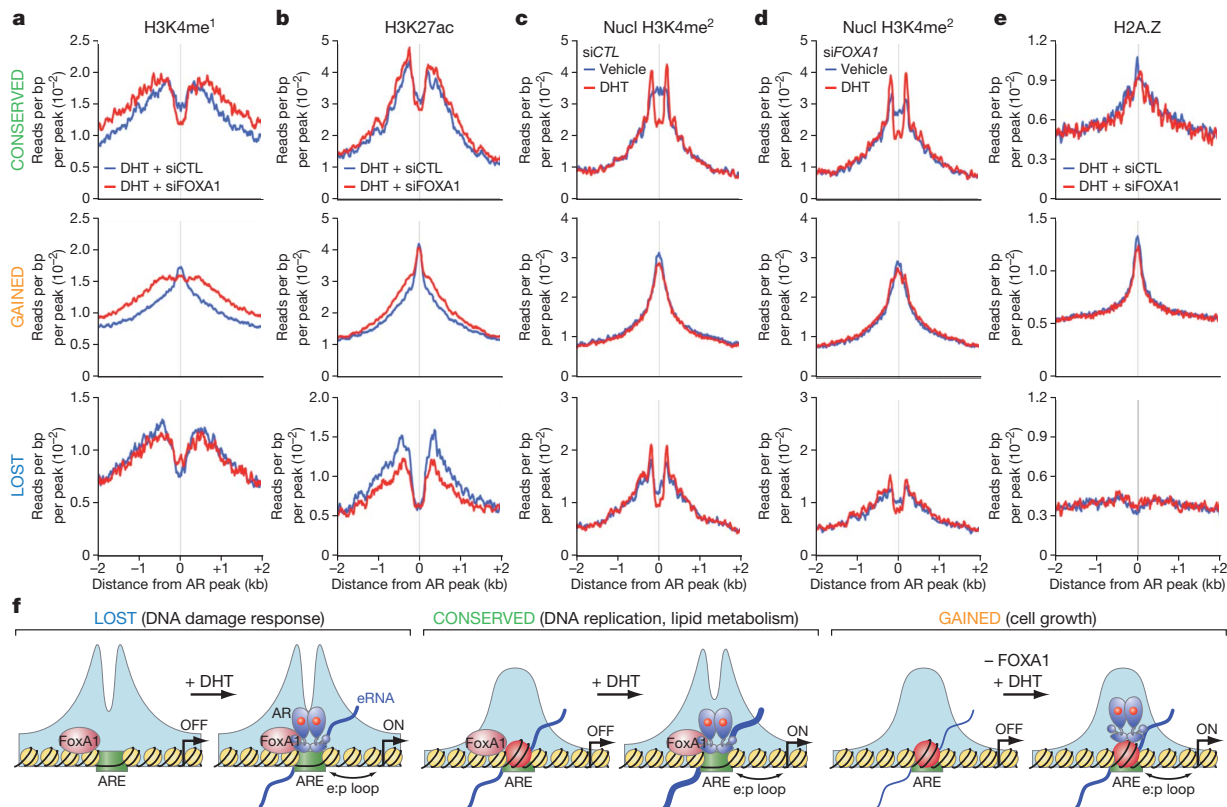


Figure 4 | Distinct classes of AR enhancers in the human genome.

a, b, Profiles of H3K4me1 (**a**) and H3K27ac (**b**) associated with the lost (bottom panels), conserved (top panels) and gained (middle panels) AR programs in DHT-treated LNCaP cells in response to *FOXA1* knockdown. **c, d,** Profiles of H3K4me2 around AR binding loci at the nucleosomal resolution in response to DHT stimulation in control siRNA-treated (**c**) or *FOXA1* siRNA-treated (**d**) LNCaP cells. **e,** Profiles of the histone variant H2A.Z on the three different AR programs. **f,** Model for FoxA1-mediated AR targeting and reprogramming in LNCaP cells. In class I (the lost AR program), FoxA1 licenses liganded AR to

bind to ARE in relatively nucleosome-free regions. AR binding does not induce nucleosome remodelling in this class of enhancers. In class II (the conserved AR program), AR binds independently of FoxA1 to ARE, inducing nucleosome remodelling. In class III (the gained AR program), FoxA1 restricts AR binding, despite the presence of strong AREs. Although pre-established, these gained loci exhibit a strong central nucleosomes and are associated with H2A.Z, which is not affected by AR binding. *FOXA1* knockdown converted these sites to androgen-responsive sites. In all these three classes, eRNAs were generated or increased after AR binding. e:p, enhancer:promoter.

is largely independent of FoxA1 (Fig. 4c, d, top panels). The gained program showed a strong H3K4me2-marked central nucleosome both before and after AR binding (Fig. 4c, d, middle panel). Thus, this gained AR program represents a new type of enhancer topography that requires no nucleosome remodelling for enhancer recognition and subsequent enhancer–promoter interactions. H2A.Z is prevalently associated with the gained AR program, modestly with the conserved AR program and absent in the lost AR program (Fig. 4e). Together, these findings establish distinct chromatin structures underlying functionally distinct classes of AR enhancer.

In summary, our findings imply a general principle for establishing cell-type-specific transcription programs. Cell-lineage-specific factors (such as FoxA1) coupled with other general transcriptional factors ‘create’ a cell-type-specific enhancer network, allowing other regulated factors (such as AR) to ‘activate’ these pre-established enhancers (Fig. 4f). The enhancer activation process is tightly linked to eRNA production, which appear to serve as a more robust indicator of enhancer activities than any enhancer-bound transcription activators or chromatin marks. On the current biology model, AR reprogramming dramatically altered the androgen-responsive pathway, which, according to GO analysis (Supplementary Fig. 22 and Fig. 23), may contribute to enhanced cell growth and the establishment of an appropriate microenvironment in advanced prostate cancer^{26–28}. Together, these findings provide a conceptual framework to understand complex gene-expression switching events, as occurs during disease progression and development.

METHODS SUMMARY

Experiments were performed on LNCaP cells, LNCaP-AR cells (gift of C. Sawyers) and HEK293 cells. ChIPs were done as previously described²⁹ and GRO was performed as described^{6,30}. Control siRNA was purchased from Qiagen (1027280). FOXA1 siRNA 1 (M-010319) and 2 (sense 5'-GAGAGAAAAAUCAACAGC-3'; antisense 5'-GCGUUGAUUUUUUCUCUC-3')⁹ were purchased from or synthesized by Dharmacon.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 July 2010; accepted 15 March 2011.

Published online 15 May 2011.

- Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
- Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genet.* **39**, 311–318 (2007).
- Pennacchio, L. A. *et al.* *In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499–502 (2006).
- Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
- Kim, T. K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
- Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
- Lee, T. I. & Young, R. A. Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* **34**, 77–137 (2000).
- Rosenfeld, M. G., Lunyak, V. V. & Glass, C. K. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev.* **20**, 1405–1428 (2006).
- Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
- Gao, N. *et al.* Forkhead box A1 regulates prostate ductal morphogenesis and promotes epithelial cell maturation. *Development* **132**, 3431–3443 (2005).
- Mirosevich, J., Gao, N. & Matusik, R. J. Expression of Foxa transcription factors in the developing and adult murine prostate. *Prostate* **62**, 339–352 (2005).

- Kang, Z., Janne, O. A. & Palvimo, J. J. Coregulator recruitment and histone modifications in transcriptional regulation by the androgen receptor. *Mol. Endocrinol.* **18**, 2633–2648 (2004).
- Shim, E. Y., Woodcock, C. & Zaret, K. S. Nucleosome positioning by the winged helix transcription factor HNF3. *Genes Dev.* **12**, 5–10 (1998).
- Holmqvist, P. H., Belikov, S., Zaret, K. S. & Wrangé, O. FoxA1 binding to the MMTV LTR modulates chromatin structure and transcription. *Exp. Cell Res.* **304**, 593–603 (2005).
- Crowe, A. J. *et al.* Hepatocyte nuclear factor 3 relieves chromatin-mediated repression of the α -fetoprotein gene. *J. Biol. Chem.* **274**, 25113–25120 (1999).
- Cirillo, L. A. *et al.* Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**, 279–289 (2002).
- Wang, Q. *et al.* Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. *Cell* **138**, 245–256 (2009).
- Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D. & Carroll, J. S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nature Genet.* **43**, 27–33 (2011).
- Thomsen, M. K. *et al.* SOX9 elevation in the prostate promotes proliferation and cooperates with PTEN loss to drive tumor formation. *Cancer Res.* **70**, 979–987 (2010).
- Wang, H. *et al.* SOX9 is expressed in human fetal prostate epithelium and enhances prostate cancer invasion. *Cancer Res.* **68**, 1625–1630 (2008).
- Ogawa, S. *et al.* Molecular determinants of crosstalk between nuclear receptors and toll-like receptors. *Cell* **122**, 707–721 (2005).
- Lee, H. J., Hwang, M., Chattopadhyay, S., Choi, H. S. & Lee, K. Hepatocyte nuclear factor-3 α (HNF-3 α) negatively regulates androgen receptor transactivation in prostate cancer cells. *Biochem. Biophys. Res. Commun.* **367**, 481–486 (2008).
- Gao, N. *et al.* The role of hepatocyte nuclear factor-3 α (Forkhead Box A1) and androgen receptor in transcriptional regulation of prostatic genes. *Mol. Endocrinol.* **17**, 1484–1507 (2003).
- Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
- He, H. H. *et al.* Nucleosome dynamics define transcriptional enhancers. *Nature Genet.* **42**, 343–347 (2010).
- Laitinen, S., Karhu, R., Sawyers, C. L., Vessella, R. L. & Visakorpi, T. Chromosomal aberrations in prostate cancer xenografts detected by comparative genomic hybridization. *Genes Chromosom. Cancer* **35**, 66–73 (2002).
- Legrier, M. E. *et al.* Hormone escape is associated with genomic instability in a human prostate cancer model. *Int. J. Cancer* **124**, 1103–1111 (2009).
- Wolf, M. *et al.* High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression. *Neoplasia* **6**, 240–247 (2004).
- Garcia-Bassets, I. *et al.* Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell* **128**, 505–518 (2007).
- Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to H.-J. Kung for providing the PC3-AR cell lines, C. Sawyers for providing the LNCaP-AR cell line, and J. Hightower, D. Benson and M. Fisher for assistance with figure and manuscript preparation. This work was supported by grants from US Army Medical Research and Materiel Command (grant W81XWH-05-1-0100) to I.G.-B., the National Institutes of Health (DK01847, DK37949, DK074868, NS34934), Department of Defense and the National Cancer Institute (CA97134) to M.G.R.; the National Institutes of Health (GM049369 and HG004659) to X.-D.F.; and the Prostate Cancer Foundation to M.G.R. and X.-D.F. M.G.R. is an investigator of the Howard Hughes Medical Institute.

Author Contributions D.W., I.G.-B., M.G.R. and X.-D.F. designed the experiments; D.W. and I.G.B. performed most of them. C.B. performed all computational analyses. W.L. and M.K. performed GRO-seq; X.S. performed the 3C assay; Y.Z. analysed human samples; and J.Q. performed proliferation assays. K.A.O. and W.L. generated AR constructs. I.G.-B., D.W., M.G.R. and X.-D.F. wrote the manuscript with contributions from C.B. and C.K.G.

Author Information High-throughput data are deposited in Gene Expression Omnibus under accession number GSE27823 for all ChIP-seq and GRO-seq experiments and GSE27682 for microarray data. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to X.-D.F. (xdfu@ucsd.edu), M.G.R. (mrosenfeld@ucsd.edu) or I.G.-B. (ibassets@ucsd.edu).

METHODS

Antibodies. Specific antibodies were purchased from the following commercial sources: anti-FoxA1 (ab5089), anti-H3K4me1 (ab8899), anti-H3K27ac (ab4729), anti-H3K36me3 (ab9050) and H2A.Z (ab4174) from Abcam; anti-AR (N-20), anti-FoxA1 (C-20) and p300 (C-20, gift of B. Ren) from Santa Cruz Biotechnology; anti-H3K4me2 (07-030), anti-H3K4me3 (07-473), anti-H4K5ac (07-327) and anti-H3K27me3 (07-449) from Upstate Biotechnology; anti-Med12 (A300-774A) from Bethyl Laboratories; and anti-beta Actin (AC74) from Sigma.

siRNA transfection. One day before transfection, LNCaP cells were seeded in RPMI 1640 medium with 10% FBS. Six hours after siRNA transfection (20 pmol ml⁻¹) with Lipofectamine 2000 (Invitrogen), cells were washed twice with PBS and then maintained in hormone-depleted phenol-free RPMI 1640 media. For gene expression profiling and western blotting, cells were cultured for 3 days after transfection and then treated with DHT for 20 h; for ChIP-qPCR and ChIP-seq, cells were cultured for 4 days after transfection and then treated with DHT for 1 h. ChIP-seq analysis at the nucleosome resolution was based on cells treated with DHT for 4 h.

3C assay. Cells were crosslinked with 1% formaldehyde for 20 min at room temperature and processed according to the standard 3C protocol³¹. For the study on the *FASN* locus, fixed chromatin from 5 × 10⁶ cells was digested with 400 units of BglII and EcoRI (NEB). For the *NDRG1* locus, fixed chromatin from 5 × 10⁶ cells was digested with 400 units of HindIII (NEB). Ligation was done with 800 units of T4 DNA ligase (NEB) for 4 h. The 3C product was quantified by qPCR after diluting the template tenfold compared with purified genomic DNA of known concentration. For each semi-quantitative PCR, the amount of template was titrated to determine the linear range in which the PCR product was amplified. PCR primers were designed next to BglII and HindIII restriction sites, respectively, for the *FASN* (all in minus strand) promoter (5'-AAGCTGTGAGTCAGCATGGTAG-3') and three upstream sites (-38kb, 5'-TGTCTCTGATGTGTCTGCTTAGAG-3'; -45kb, 5'-AATCCTGCTCAGGAATCTGTATGT-3'; -54kb, 5'-GGACACTACTGCTTTTCTGTG-3') and for the *NDRG1* (all in plus strand) promoter (5'-ATAGGTTCTGCCTTATTAGGG-3') and three upstream sites (-42kb, 5'-ATAGAGTTAGAGAAACGGAGGCAGT-3'; -56kb: 5'-GCCGTGAAGATAAACAAAGATGAG-3'; -62kb: 5'-ACACATTTTGTCCCAGTGCAG-3').

Co-IP and western blotting analysis. HEK293 cells were seeded for 1 day, transfected with the expression plasmids expressing wild-type, mutant AR and FoxA1 using Lipofectamine2000 (Invitrogen) and then changed to hormone-depleted, phenol-free DMEM medium. One day after plasmid transfection, cells were treated with 100 nM DHT for another day. Cells were washed by cold PBS twice and treated with 1 ml of lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1% NP-40) supplemented with a cocktail of proteinase inhibitors (Sigma) for 5 min at 4 °C. Lysed cells were collected, rotated for 1 h at 4 °C and cell debris removed by centrifugation at 18,000g for 30 min in a cold room. The supernatant was incubated with anti-AR, anti-FoxA1 or immunoglobulin-G overnight at 4 °C followed by the addition of 50 µl of 50% protein G beads to each tube. After rotating for another 2 h at 4 °C, the beads were washed five times with the lysis buffer, twice with cold PBS, and boiled for 6 min in 40 µl of 2 × SDS loading buffer. Western blotting analysis was performed with anti-AR or anti-FoxA1.

Luciferase reporter assay. PC3-AR and HEK293 cells were seeded into 24-well plates in hormone-depleted and phenol-free RPMI 1640 medium and DMEM 1 day before transfection. Transfection was performed according to the manufacturers' recommendations (DOTAP Liposomal Transfection Reagent from Roche or Lipofectamin 2000 from Invitrogen). One day after transfection, these cells were treated with DHT for an additional day. After washing with cold PBS twice, cells were treated with the lysis buffer (Promega) and the Luciferase signal was recorded.

Cell proliferation assay. The assay was based on the published protocol³². Briefly, LNCaP cells were transfected with control siRNA and FoxA1 siRNA (sequences listed in Methods Summary) and cultured in hormone-depleted medium for 3 days. The cells were treated with different amount of DHT for another day. After the treatment, cells were washed by PBS, fixed by 70% EtOH and stored at -20 °C for at least 2 h. Before analysis, cells were washed with cold PBS, resuspended at the propidium-iodide/Triton X-100 staining solution and incubated at 37 °C for 15 min. After removing cell clumps, stained cells were sorted on a Beckman FAScan, and the percentage of S-phase cells was calculated.

ChIP and ChIP-seq analyses. ChIP was as previously described²⁹. Briefly, approximately 10⁷ treated cells were crosslinked with 1% formaldehyde at room temperature for 15 min. After sonication, the soluble chromatin was incubated with 1–5 µg of antibody. Specific immunocomplexes were precipitated with Protein A/G beads (Sigma-Aldrich). Complexes were washed, DNA extracted and purified by QIAquick Spin columns (Qiagen). Extracted DNA (1 µl from 60 µl) was used for qPCR with the specific PCR primers listed in Supplementary

Fig. 24, each of which was designed surrounding a specific region of 150–250 base pairs (bp) on target DNA. PCR products were detected with SYBR Green on a MX3000P System (Stratagene) and the percentage of immunoprecipitated chromatin was calculated from ΔC_t relative to immunoglobulin-G control after normalizing against input chromatin. For ChIP-seq, extracted DNA was ligated to specific adaptors followed by deep sequencing in the Illumina GAI system according to the manufacturer's instructions. The first 25 bp for each sequence tag returned by the Illumina Pipeline was aligned to the hg18 assembly (National Center for Biotechnology Information, build 36.1) using Bowtie, allowing up to two mismatches. Only tags uniquely mapped to the genome were used for further analysis. The data were visualized by preparing custom tracks for the University of California, Santa Cruz, genome browser using HOMER³³ (<http://biowhat.ucsd.edu/homer>). The total number of mappable reads was normalized to 10⁷ for each experiment presented in this study. ChIP-seq at nucleosome resolution was performed as previously reported³⁴. A summary of ChIP experiments is provided in Supplementary Fig. 25.

Identification of ChIP-seq peaks. The identification of ChIP-seq peaks (bound regions) was performed using HOMER (<http://biowhat.ucsd.edu/homer>). For transcription factors, peaks were identified by searching locations of high read density using a 200-bp sliding window. Regions of maximal density exceeding a given threshold were called as peaks, and we required adjacent peaks to be at least 500 bp away to avoid redundant detection. Only one tag from each unique position was considered to avoid clonal artefacts from the sequencing. The threshold for the number of tags that determined a valid peak was selected at a false discovery rate of 0.001 determined by peak finding using randomized tag positions in a genome with an effective size of 2 × 10⁹ bp. We also required peaks to have at least fourfold more tags (normalized to total count) than input control samples. In addition, we required fourfold more tags relative to the local background region (10 kb) to avoid identifying regions with genomic duplications or non-localized binding.

The peak finding procedure was modified to identify regions harbouring specific histone modifications, as these experiments tend to yield broad areas of enrichment over several hundreds or thousands of base pairs. Seed regions were initially found using a peak size of 500 bp at the false discovery rate of 0.001 to identify enriched loci. Enriched loci found within 1 kb of one another were then merged to yield variable-length regions. Transcription factor peaks and histone modification regions were associated with gene products by identifying the nearest RefSeq TSS. Annotated positions for promoters, exons, introns and other features were based on RefSeq transcripts and repeat annotations from University of California, Santa Cruz. Peaks from separate experiments were considered equivalent/co-bound if their peak centres were located within 200 bp of each other. Read density heat maps were created by first using HOMER to generate read densities and then visualized using Java TreeView (<http://jtreeview.sourceforge.net>).

HOMER for de novo motif discovery and known motif enrichment. Motif discovery was performed using a comparative algorithm similar to those previously described³⁵. An in-depth description will be published elsewhere (Benner *et al.*, in preparation). Motif finding for transcription factors was performed on sequence from ±100 bp relative to the peak centre, whereas motif finding for histone modification regions was performed on sequence from ±500 bp relative to the region centre. Briefly, sequences were divided into target and background sets for each application of the algorithm. Background sequences were then selectively weighted to equalize the distributions of G + C content in target and background sequences to avoid comparing sequences of different general sequence content. Motifs of length 8, 10, 12, 14, 16 and 18 bp were identified separately by first exhaustively screening all oligonucleotides for enrichment in the target set compared with the background set using the cumulative hypergeometric distribution to score enrichment. Up to three mismatches were allowed in each oligonucleotide sequence to increase the sensitivity of the method. The top 200 oligonucleotides of each length with the lowest *P* values were then converted into probability matrices and heuristically optimized to maximize hypergeometric enrichment of each motif in the given data set. As optimized motifs were found they were removed from the data set to facilitate the identification of additional motifs in subsequent rounds. HOMER also screens the enrichment of known motifs previously identified through the analysis of published ChIP-ChIP and ChIP-Seq data sets by calculating the known motifs' hypergeometric enrichment in the same set of G + C normalized sequences used for *de novo* analysis. Sequence logos were generated using WebLogo (<http://weblogo.berkeley.edu>). Motif enrichment heatmaps and dendrograms were created by clustering hypergeometric log *P* values using Cluster (<http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm#ctv>) and Java TreeView (<http://jtreeview.sourceforge.net>).

GRO-seq. Global run-on⁶ and library preparation for sequencing³⁰ were done as described. Briefly, four 10-cm plates of confluent LNCaP cells per treatment were washed three times with cold PBS buffer. Cells were then swelled in swelling buffer (10 mM Tris pH 7.5, 2 mM MgCl₂, 3 mM CaCl₂) for 5 min on ice. Harvested cells

were re-suspended in 1 ml of the lysis buffer (swelling buffer with 0.5% IGEPAL and 10% glycerol) with gentle vortex and brought to 10 ml with the same buffer for extraction of nuclei. Nuclei were washed with 10 ml of lysis buffer and re-suspended in 1 ml of freezing buffer (50 mM Tris pH 8.3, 40% glycerol, 5 mM MgCl₂, 0.1 mM EDTA), pelleted down again and finally re-suspended in 100 µl of freezing buffer.

For run-on assay, re-suspended nuclei were mixed with an equal volume of reaction buffer (10 mM Tris pH 8.0, 5 mM MgCl₂, 1 mM DTT, 300 mM KCl, 20 units of SUPERase-In, 1% Sarkosyl, 500 µM ATP, GTP and Br-UTP, 2 µM CTP) and incubated for 5 min at 30 °C. Nuclei RNA were extracted with TRIzol LS reagent (Invitrogen) according to the manufacturer's instructions. RNA was then re-suspended in 20 µl of DEPC-water and subjected to base hydrolysis by addition of 5 µl of 1 M NaOH and incubated on ice for 40 min. Then, 25 µl of 1 M Tris pH 6.8 was added to neutralize the reaction. RNA was purified through a p-30 RNase-free spin column (BioRad), according to the manufacturer's instructions and treated with 6.7 µl of DNase buffer and 10 µl of RQ1 RNase-free DNase (Promega), and purified again through a p-30 column. A volume of 8.5 µl 10 × antartctic phosphatase buffer, 1 µl of SUPERase-In and 5 µl of antartctic phosphatase was added to the run-on RNA and treated for 1 h at 37 °C. Before proceeding to immunoprecipitation, RNA was heated to 65 °C for 5 min and kept on ice.

Anti-BrdU argarose beads (Santa Cruz Biotech) were blocked in blocking buffer (0.5 × SSPE, 1 mM EDTA, 0.05% Tween-20, 0.1% PVP, and 1 mg ml⁻¹ BSA) for 1 h at 4 °C. Heated run-on RNA (~85 µl) was added to 60-µl beads in 500 µl binding buffer (0.5 × SSPE, 1 mM EDTA, 0.05% Tween-20) and allowed to bind for 1 h at 4 °C with rotation. After binding, beads were washed once in low salt buffer (0.2 × SSPE, 1 mM EDTA, 0.05% Tween-20), twice in high salt buffer (0.5% SSPE, 1 mM EDTA, 0.05% Tween-20, 150 mM NaCl) and twice in TET buffer (TE pH 7.4, 0.05% Tween-20). BrU-incorporated RNA was eluted with 4 × 125 µl elution buffer (20 mM DTT, 300 mM NaCl, 5 mM Tris pH 7.5, 1 mM EDTA and 0.1% SDS). RNA was then extracted with acidic phenol/chloroform once, chloroform once and precipitated with ethanol overnight. The precipitated RNA was re-suspended in 50 µl reaction (45 µl of DEPC water, 5.2 µl of T4 PNK buffer, 1 µl of SUPERase-In and 1 µl of T4 PNK (NEB)) and incubated at 37 °C for 1 h. The RNA was extracted and precipitated again as above.

Complementary DNA (cDNA) synthesis was performed basically as described³⁰ with minor modifications. First, RNA fragments were subjected to poly-A tailing reaction in 8.0 µl volume containing 0.8 µl poly-A polymerase buffer, 1 µl 1 mM ATP, 0.5 µl SUPERase-In and 0.75 µl poly-A polymerase (NEB). The reaction was performed for 30 min at 37 °C. Subsequently, reverse transcription was performed using oNTI223 primer (5'-pGATCGTCCGACTGTAGAAGCTCT;CAAGCAGAAGACGGCATAACGATTTTTTTTTTTTTTTTTTTVN-3') where the p indicates 5' phosphorylation, 'i' indicates the abasic dSpacer furan and VN indicates degenerate nucleotides.

Tailed RNA (8.0 µl) was mixed with 1 µl dNTP (10 mM each) and 2.5 µl 12.5 µM oNTI223, heated for 3 min at 75 °C and chilled briefly on ice. Then, 0.5 µl SUPERase-In, 3 µl 0.1 M DTT, 2 µl 25 mM MgCl₂, 2 µl 10 × reverse transcription buffer and 1 µl superscript III reverse transcriptase (Invitrogen) was added to the tube. The tube was incubated for 30 min at 48 °C. After that, 4 µl of Exonuclease I (Fermentas) was added into the reaction and the tube was incubated for 1 h at 37 °C to eliminate extra oNTI223. Then RNA was eliminated by adding 1.8 µl 1 M NaOH and incubated for 20 min at 98 °C. The reaction was neutralized with 1.8 µl of 1 M HCl. After running on a 10% polyacrylamide TBE-urea gel, the extended first-strand cDNA product was excised and recovered by soaking the grinded gel in DNA gel elution buffer (TE with 0.1% Tween-20 and 150 mM NaCl) overnight and then precipitated with ethanol.

Circularization of first-strand cDNA was performed by re-suspending cDNA in 9.5 µl reaction solution (7.5 µl water, 1 µl CircLigase buffer, 0.5 µl 1 mM ATP and 0.5 µl 50 mM MnCl₂) and then adding 0.5 µl CircLigase (Epicentre). The reaction went for 1 h at 60 °C and then was heat-inactivated for 20 min at 80 °C. Circularized single-stranded DNA (ssDNA) was religated by adding 3.8 µl of 4 × religation supplement (100 mM KCl, 2 mM DTT) followed by 1.5 µl of ApeI (15u, NEB). The reaction was incubated for 1 h at 37 °C. Religated ssDNA was separated in a 10% polyacrylamide TBE-urea gel (Invitrogen) as described above. The religated product band was excised (~120–300 bp) and the DNA was recovered as described above.

The ssDNA template was amplified by PCR using the Phusion High-Fidelity enzyme (NEB) according to the manufacturer's instructions. The oligonucleotide primers oNTI200 (5'-CAAGCAGAAGACGGCATA-3') and oNTI201 (5'-AATGATACGCGCACCACGACAGGTTACAGATTCTACAGTCCGACG-3') were used to generate DNA for sequencing. PCR was performed with an initial 30-s denaturation at 98 °C, followed by 13 cycles of 10-s denaturation at 98 °C, 15-s annealing at 60 °C and 15-s extension at 72 °C. The PCR product was run on a non-denaturing 8% polyacrylamide TBE gel and recovered as mentioned before.

DNA was then sequenced on the Illumina Genome Analyser II according to the manufacturer's instructions with small RNA sequencing primer 5'-CGACAGGTTTCAGAGTTCTACAGTCCGACGATC-3'.

De novo identification of GRO-Seq transcripts. To identify transcription units in an unbiased manner, GRO-Seq read densities were analysed to classify genomic regions into contiguous transcripts using HOMER. GRO-Seq read densities were initially normalized using the GC content of individual reads to remove any systematic bias introduced by overall GC content variation between read libraries. To maximize read depth for transcript identification, all GRO-Seq libraries were merged to perform the initial transcript discovery, and later considered separately to identify regulated transcripts. For each strand of each chromosome, GRO-Seq read densities were calculated using a sliding window of 250 bp. Regions for which GRO-Seq reads could not be uniquely mapped (that is, repeats) were first identified and then read densities in these regions were estimated using upstream read densities from mappable regions to avoid ending predicted transcripts prematurely. Transcript initiation sites were identified as regions where the GRO-Seq read density increased threefold relative to the previous 1 kb region. Transcript termination sites were defined by either a reduction in reads below 10% of the start of the transcript or when another transcript's start site occurred on the same strand. Single spikes in read density covering a span less than 250 bp were considered artefacts and discarded. Identified transcripts were strand-specifically compared with RefSeq transcripts by looking for overlap in the transcribed region. Transcripts were defined as putative eRNAs if their TSS was located distal to RefSeq TSS (>3 kb) and were associated with H3K4me1 regions. To identify differentially regulated transcripts, strand-specific read counts from each GRO-Seq experiment were determined for each transcript using HOMER³³. EdgeR (<http://www.bioconductor.org/packages/release/bioc/html/edgeR.html>) was then used to calculate differential expression (>1.5-fold, <0.01 false discovery rate).

Microarray and reverse-transcription qPCR analyses of gene expression. Total RNA was isolated with the RNeasy Mini Kit (Qiagen) and treated by RNase-free DNase I. For PCR with reverse transcription, first-strand cDNA synthesis from total RNA was performed with the Superscript III cDNA Synthesis System (Invitrogen). Microarray analysis was performed on Human V2 Chips (Illumina). The published gene expression profiling data GDS2545 (refs 36, 37) and GDS1439 (ref. 38) were extracted from the National Center for Biotechnology Information, normalized and *P* values calculated by two-tailed *t*-test. For validation by PCR with reverse transcription, cDNA was analysed with SYBR Green (Stratagene) on the Mx300P System (Stratagene). The relative messenger RNA level was calculated by comparing with non-treatment control, after normalization with *GAPDH* or *ACTB* messenger RNA. The primers for RT-qPCR (5'–3') were as follows: *ACTB*-5, CGTCCAGTGTGTGACGATG; *ACTB*-3, GCCGTCTTCCCTCCATC; *GAPDH*-5, GTTTTCTAGACGGCAGGTGTCAGG; *GAPDH*-3, AACATCATCCTGCCTCTACTGG; *KLK3*-5, TGTGTGCGCAAGTTCAAC; *KLK3*-3, GGTTCACTGCCCATGAC; *RASSF3*-5, GACGCCGAGGACTTCTTCTT; *RASSF3*-3, TGCTGAGGTAAGTGTGGGTTT; *SOX9*-5, GACTCGCCACACTCCTCCTC; *SOX9*-3, AAGTCGATAGGGGGCTGTCT; *IL6R*-5, GAGATTCTGCAAAATGCGACA; *IL6R*-3, GTGGGAGATGAGAGGAACA; *DNM2*-5, TGTTTGCCAAACAGTGACCTC; *DNM2*-3, CCCAGACCATGAAGCTCCT.

Survival analysis. Two independent sets of gene expression data were used to check the association between FoxA1 and clinical outcome of patients by Kaplan–Meier analysis. One data set came from 78 patients with prostate tumours (age <70)³⁹, the other from 131 patients⁴⁰. Significant association with outcome was determined by log-rank test for survival. Hazard ratios were calculated by the Cox proportional model. All statistics were analysed with the statistical software R (version 2.6.2), available from the R Project for Statistical Computing website (<http://www.r-project.org>). The cut-off was determined so that the log-rank *P* value was the smallest one in the cut-offs that went through the 5th–95th percentiles of signals.

- Hu, Q. *et al.* Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc. Natl Acad. Sci. USA* **105**, 19199–19204 (2008).
- Holbro, T. *et al.* The ErbB2/ErbB3 heterodimer functions as an oncogenic unit: ErbB2 requires ErbB3 to derive breast tumor cell proliferation. *Proc. Natl Acad. Sci. USA* **100**, 8933–8938 (2003).
- Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
- Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
- Linhardt, C., Halperin, Y. & Shamir, R. Transcription factor and microRNA motif discovery: the Amadeus platform and a compendium of metazoan target sets. *Genome Res.* **18**, 1180–1189 (2008).
- Chandran, U. R. *et al.* Gene expression profiles of prostate cancer reveal involvement of multiple molecular pathways in the metastatic process. *BMC Cancer* **7**, 64 (2007).

37. Yu, Y. P. *et al.* Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. *J. Clin. Oncol.* **22**, 2790–2799 (2004).
38. Varambally, S. *et al.* Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell* **8**, 393–406 (2005).
39. Glinsky, G. V., Glinskii, A. B., Stephenson, A. J., Hoffman, R. M. & Gerald, W. L. Gene expression profiling predicts clinical outcome of prostate cancer. *J. Clin. Invest.* **113**, 913–923 (2004).
40. Taylor, B. S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**, 11–22 (2010).

Interannual atmospheric variability forced by the deep equatorial Atlantic Ocean

Peter Brandt¹, Andreas Funk¹, Verena Hormann^{1†}, Marcus Dengler¹, Richard J. Greatbatch¹ & John M. Toole²

Climate variability in the tropical Atlantic Ocean is determined by large-scale ocean–atmosphere interactions, which particularly affect deep atmospheric convection over the ocean and surrounding continents¹. Apart from influences from the Pacific El Niño/Southern Oscillation² and the North Atlantic Oscillation³, the tropical Atlantic variability is thought to be dominated by two distinct ocean–atmosphere coupled modes of variability that are characterized by meridional^{4,5} and zonal^{6,7} sea-surface-temperature gradients and are mainly active on decadal and interannual timescales, respectively^{8,9}. Here we report evidence that the intrinsic ocean dynamics of the deep equatorial Atlantic can also affect sea surface temperature, wind and rainfall in the tropical Atlantic region and constitutes a 4.5-yr climate cycle. Specifically, vertically alternating deep zonal jets of short vertical wavelength with a period of about 4.5 yr and amplitudes of more than 10 cm s^{-1} are observed, in the deep Atlantic, to propagate their energy upwards, towards the surface^{10,11}. They are linked, at the sea surface, to equatorial zonal current anomalies and eastern Atlantic temperature anomalies that have amplitudes of about 6 cm s^{-1} and 0.4°C , respectively, and are associated with distinct wind and rainfall patterns. Although deep jets are also observed in the Pacific¹² and Indian¹³ oceans, only the Atlantic deep jets seem to oscillate on interannual timescales. Our knowledge of the persistence and regularity of these jets is limited by the availability of high-quality data. Despite this caveat, the oscillatory behaviour can still be used to improve predictions of sea surface temperature in the tropical Atlantic. Deep-jet generation and upward energy transmission through the Equatorial Undercurrent warrant further theoretical study.

Tropical Atlantic variability, which modulates the seasonal migration of the intertropical convergence zone, is dominated by two modes of behaviour^{8,9}. The meridional mode, peaking during boreal spring, is characterized by a north–south sea-surface-temperature (SST) gradient that drives cross-equatorial wind anomalies from the cold hemisphere to the warm^{4,5}. The zonal mode is characterized by an east–west SST gradient along the Equator and is associated with marked zonal wind anomalies^{6,7}. It is most pronounced during boreal summer when the seasonal maximum in equatorial upwelling leads to the development of the eastern Atlantic SST cold tongue. The zonal mode is often referred to as the Atlantic counterpart to the Pacific El Niño. The period of zonal-mode-like oscillations estimated from observations, models and theory ranges from 19 months to 4 years^{6,14–16}. However, aspects of the intrinsic ocean dynamics, such as year-to-year variations in the strength of tropical instability waves, are similarly identified as causes of interannual SST variability¹⁷ and may themselves be able to force variability in the atmosphere.

During the past 10–20 yr, the eastern equatorial Atlantic SST, represented by the ATL3 index (that is, the average SST anomaly inside the box shown in Fig. 1a), has shown pronounced variability on interannual timescales, dominated by the period range of 4–5 yr; maximum explained variance of different ocean parameters is found at a period of 1,670 d (Supplementary Fig. 1). The associated harmonic amplitude of local

SST fluctuations, which is $0.29 \pm 0.08^\circ\text{C}$ averaged over the ATL3 region, is generally high in the eastern equatorial Atlantic, with local amplitudes of up to 0.4°C (Fig. 1a and Supplementary Fig. 2). The regression of surface winds and rainfall on the 1,670-d SST harmonic reveals that anomalous westerlies along the Equator, convergent meridional wind anomalies particularly in the western tropical Atlantic, and positive rainfall anomalies in a wide belt around the Equator are associated with positive SST anomalies.

A 1,670-d cycle is also found in the surface geostrophic zonal velocity anomaly at the Equator and is again the dominant interannual variability, with a harmonic amplitude of $5.9 \pm 1.9 \text{ cm s}^{-1}$. Phases of eastward surface flow coincide with SST warm phases in the eastern equatorial Atlantic (Fig. 1b). Whereas the 1,670-d period stands out as the dominant interannual variability timescale of the equatorial zonal surface flow, this is not the case for the wind forcing, which instead shows more irregular fluctuations during the analysed time interval (NCEP/NCAR reanalysis wind data). Such a dominant signal in the ocean seems to contradict early model results, in which the equatorial ocean response to wind forcing with periods longer than about 150 d was found to be a succession of equilibrium responses with the strength of the flow independent of the forcing period¹⁸. As we show below, variability in the 4–5-yr period band is a ubiquitous feature of the equatorial Atlantic and, furthermore, is associated with upward propagation of energy in the ocean. We propose that the variability in the equatorial zonal surface flow is not due to wind forcing with the same period but rather is a mode internal to the ocean, with its origin in the abyss (perhaps as deep as several thousand metres). If this is indeed the case, then the observed atmospheric variability in the 4–5-yr period band in the equatorial Atlantic can be interpreted as a consequence of internal ocean dynamics.

Analysis of zonal velocities at 1,000-m depth as observed by Argo floats¹⁹ reveals periodic behaviour similar to that of the SST and surface geostrophic zonal velocity anomalies (Fig. 1b). The dominant period, of 4.4 yr, in the Argo float drift data for the period 1998–2010 is in agreement with earlier estimates from moored zonal velocity observations in the depth range 600–1,800 m made during 2000–2006¹¹ (4.4 yr) and with the estimate from hydrographic observations made during 1972–1998¹⁰ (5 ± 1 yr). The deep velocity and density fluctuations have been dynamically described as a mixture of high-baroclinic-mode Kelvin and Rossby waves representing quasi-steady equatorial deep jets^{10,11}. Such vertically alternating zonal jets with vertical wavelengths between 300 and 700 m are similarly present in the Pacific^{12,20} and Indian oceans^{13,21}. In the Atlantic, a downward phase velocity of equatorial deep jets (of about 100 m yr^{-1}) is observed¹¹ that corresponds, according to linear internal wave theory, to upward energy propagation. Our moored observations reveal downward phase propagation from below the Equatorial Undercurrent (EUC) at about 200-m depth to about 2,000-m depth (Fig. 2 and Supplementary Fig. 3), suggesting a deep generation mechanism for equatorial deep jets. Observed variations in the vertical phase velocity are probably due to changes in the amplitudes of different superimposed baroclinic modes, as also indicated by changes in the vertical wavelength (Fig. 2). Theories of

¹IFM-GEOMAR, Leibniz-Institut für Meereswissenschaften an der Universität Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany. ²Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA. [†]Present address: Cooperative Institute for Marine and Atmospheric Studies, University of Miami, and National Oceanic and Atmospheric Administration/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida 33149, USA.

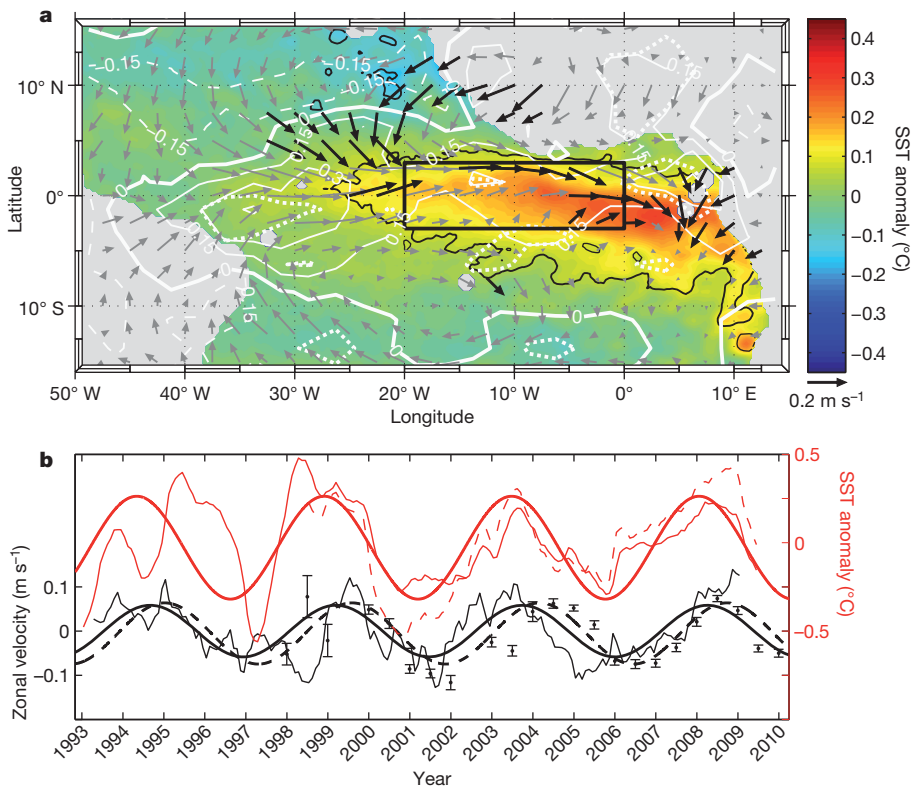


Figure 1 | Interannual variability in the tropical Atlantic associated with a 1,670-d cycle.

a, Anomalies of SST (colour scale), surface wind (arrows) and rainfall (white contours: solid, positive; dashed, negative; every 0.15 mm d^{-1}) as determined through regression on the harmonic fit of the SST anomalies (microwave optimally interpolated SST) averaged within the marked box (ATL3: 3° S – 3° N , 20° W – 0°). We mark significant correlations (95%) of harmonic fit with SSTs (black lines), winds (black arrows) and rainfall (white dotted lines). **b**, ATL3 SST anomaly (microwave optimally interpolated SST), red dashed; HadISST, red thin solid) with 1,670-d harmonic fit (red thick solid), surface zonal velocity anomaly (Equator, 35° W – 15° W ; black thin solid) with 1,670-d harmonic fit (black thick solid), and 1,000-m zonal velocity (1° S – 1° N , 35° W – 15° W ; black dots with standard errors) with 1,670-d harmonic fit (black thick dashed).

deep-jet generation involve instabilities associated with the propagation of intraseasonal mixed Rossby gravity waves^{22,23} or the Equator-crossing deep western boundary current²⁴. However, until now the proposed theories have failed to explain the observed strength and complex behaviour of the deep jets in the different ocean basins.

Propagation of deep-jet energy towards the surface is complicated by the presence of a strong, vertically-sheared mean current, the EUC, with maximum eastward velocities of more than 60 cm s^{-1} at about 80-m depth (Fig. 3a). Theoretical studies indicate that the EUC effectively modifies dispersion characteristics of Kelvin and Rossby waves²⁵. On seasonal timescales, the background flow partly inhibits

the downward propagation of high-baroclinic-mode energy, explaining the dominance of low-baroclinic-mode seasonal waves at depth. Theoretical studies of internal wave propagation motivated by observed internal wave transmissions across an atmospheric jet suggest, however, that an energy transfer across critical levels—that is, where the horizontal phase velocity equals the background mean flow—is possible²⁶.

The amplitude of the 1,670-d harmonic oscillation of zonal velocity in the upper 600 m of the water column is largest in the 300–600-m depth interval (Fig. 3b), where it explains up to 60% of the variance contained in the monthly zonal velocity anomalies (Fig. 3d). Local minima in the amplitude of the 1,670-d oscillation are indicated near

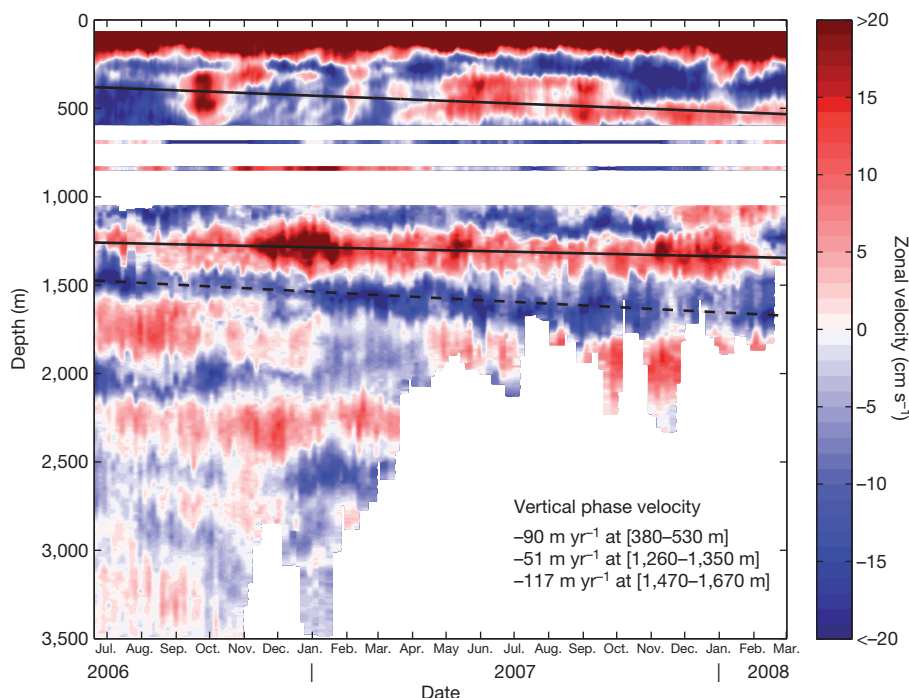


Figure 2 | Zonal velocities at the Equator, 23° W .

Velocity data above 600 m are from a moored acoustic Doppler current profiler with annual and semi-annual cycles subtracted, those between 600 and 1,000 m are from two single-point current meters, and those below 1,000 m are from a moored profiler. The white areas mark depths not sampled by the deployed instrumentation. Linearized phase lines (eastward jets, solid; westward jet, dashed) of equatorial deep jets are calculated from about 7-yr of moored current data (above 600 m) and from the presented data (below 1,000 m). Associated vertical phase velocities are given in the figure.

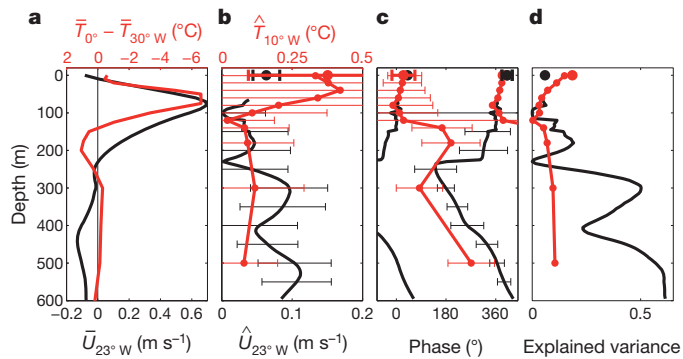


Figure 3 | Mean zonal velocity, zonal temperature gradient and harmonic analysis of 1,670-d oscillation. **a**, Moored mean zonal velocity (\bar{U}) at the Equator, 23°W (black), and climatological²⁷ mean zonal temperature (\bar{T}) difference at the Equator between 0° and 30°W (red). **b–d**, 1,670-d harmonic amplitude (**b**), phase (**c**) and explained variance (**d**) of equatorial moored zonal velocities at 23°W (black curves ($\bar{U}_{23^\circ W}$)), equatorial surface zonal velocity averaged between 35°W and 15°W (black dots), and subsurface temperatures (red curves ($\bar{T}_{10^\circ W}$) and small red dots) and microwave optimally interpolated SST (big red dot) at the Equator, 10°W. Zero phase corresponds to 1 January 1993; explained variance is calculated using monthly mean data with the mean seasonal cycle subtracted. Information on the calculation of error bars in **b** and **c** can be found in Methods.

the core and at the lower boundary of the EUC (Fig. 3b), and amplitudes of about 6 cm s^{-1} are derived at the surface. The variance explained by the 1,670-d harmonic oscillation decreases towards the surface (Fig. 3d), mainly as a result of the increasing strength of intraseasonal fluctuations. Although the vertical phase propagation is consistently downward below the EUC, the phase jumps by about 180° at the lower boundary of the EUC (Fig. 3c), approximately at the critical level for the propagation of high-baroclinic-mode equatorial Kelvin waves.

The 1,670-d fluctuations are also pronounced in subsurface temperature records. Temperatures are affected in two ways by the

presence of equatorial deep jets: isopycnal displacements associated with the deep jets will lead to temperature variations that are phase-shifted in space and time relative to the velocity anomalies, depending on the character (Rossby or Kelvin) of the wave¹⁰; and in the presence of climatological zonal temperature gradients, zonal advection associated with the jets might induce changes in the temperature fields. For example, in-phase oscillations of surface zonal velocity and near-surface temperatures (Fig. 3c) are in agreement with the propagation of equatorial Kelvin waves; that is, eastward velocities are associated with downward isopycnal (isothermal) displacements and vice versa. A deeper thermocline could, in turn, be associated with reduced downward heat transport through diapycnal mixing causing higher SSTs. In the equatorial Atlantic, the climatological²⁷ zonal temperature gradient changes sign with depth, further complicating the interpretation of the observed phase structure of the subsurface temperature variability: for example, the reversal of the zonal temperature gradient with depth in the lower part of the EUC (Fig. 3a) might be responsible for the phase shift with depth of the 1,670-d harmonic oscillation of the subsurface temperature (Fig. 3c). Although the understanding of the propagation characteristics of the jets in the presence of strong mean currents and zonal tracer gradients deserves further theoretical study, these observations suggest that equatorial deep-jet energy propagates to the surface and affects sea surface conditions.

Observations in the equatorial Atlantic reveal a similar periodic behaviour for deep-jet oscillations over different time intervals and depth ranges^{10,11}. Such consistent behaviour could arise from the development of high baroclinic basin modes²² established by the eastward and westward propagation of Kelvin and Rossby waves, respectively²⁸. In this case, vertical phase and energy propagation can occur only for quasi-resonant modes with active forcing and dissipation. The basin width of the Indian Ocean suggests a similar period for equatorial deep-jet oscillations as in the Atlantic, with rather different behaviour in the Pacific as a result of the much greater basin width. Argo float drift data from about 1,000-m depth represent a consistent data set that is available for all three oceans¹⁹. In the Atlantic, maximum

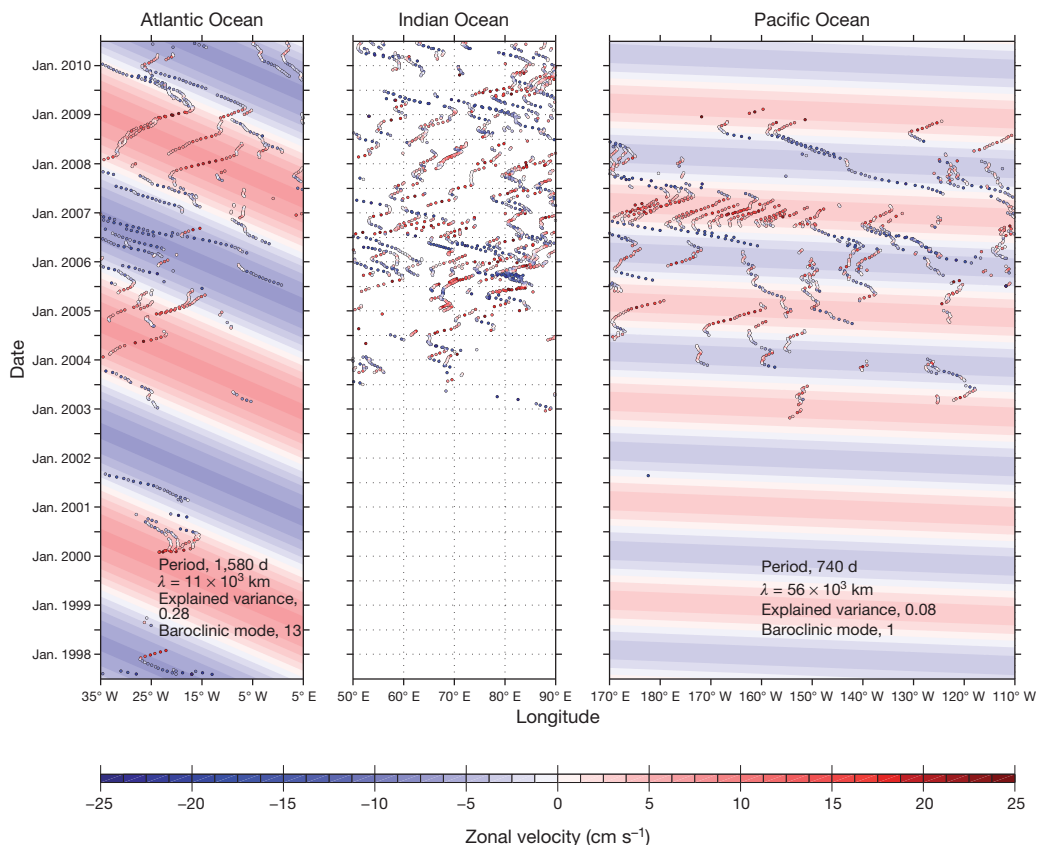


Figure 4 | Equatorial zonal velocities from 1,000-m Argo float drift data. Argo float drift data (coloured dots, colour scale) were acquired between latitudes 1°S and 1°N . The dominant interannual variability in the Atlantic and Pacific oceans obtained by maximizing explained variance using a plane-wave fit is visualized by colour shadings. Associated harmonic parameters for the Atlantic and Pacific oceans are given in the figure.

explained variance is found for westward-propagating Rossby waves of baroclinic mode 13 (corresponding to a vertical wavelength of about 600 m at 1,000-m depth) and a period of 1,580 d, corresponding to a zonal wavelength ($\lambda = 2\pi/|k|$, where k is the zonal wavenumber) of 11×10^3 km. In the Pacific, only weak signals of high-baroclinic-mode variability were extracted from the approximately 7-yr-long time series, which could be expected as estimated deep-jet oscillation periods are in the multidecadal range^{12,20}. The dominant signal there is associated with low-baroclinic-mode variability. Despite there being geometric similarities between the Indian and Atlantic oceans, during the analysed time frame the Indian Ocean Argo float velocities are characterized by incoherent signals in the interannual period range, with no preferred period (Fig. 4). From this analysis, we expect no influence of equatorial deep jets on the surface conditions in the Indian and Pacific oceans on interannual timescales.

In analysing the seasonality of the Atlantic deep-jet surface expressions, we find that the amplitude of the 1,670-d cycle of zonal velocity is seasonally independent whereas the corresponding amplitudes of the ATL3 SST anomalies at this period are instead strongest during boreal summer and November/December (Supplementary Figs 6 and 7). These periods are identified as cold seasons with shallow thermocline depths in the east and active Bjerknes positive feedback^{29,30}. During boreal spring when the tropical Atlantic is uniformly warm, the influence of the 1,670-d zonal velocity anomalies on SST is weak. Such behaviour is consistent with the equatorial zonal surface flow forced by interior ocean dynamics, whereas associated SST variations are seasonally modulated. On decadal timescales, the strength and period of the deep-jet oscillations may vary over time. The modulation could be due, for example, to a change in the dominant baroclinic mode affecting the basin mode period^{22,28}. Such behaviour is suggested by Supplementary Fig. 8, although other modes of variability, such as the Pacific El Niño/Southern Oscillation² and the North Atlantic Oscillation³, could also be influencing the time series. Despite this caveat, the surface expressions of the deep jets can clearly be used to improve the prediction of equatorial Atlantic SST, which is crucial for seasonal to interannual climate forecasting in the region⁹.

METHODS SUMMARY

We calculated surface zonal velocity anomaly at the Equator, averaged between 35° W and 15° W (Fig. 1b), by applying a second-order fit in latitude to monthly mean meridional sea level anomaly distributions between 1° N and 1° S and evaluating equatorial geostrophy using the obtained curvature. The standard error of annual mean Argo float velocities (Fig. 1b) was calculated by dividing their standard deviation by the square root of the number of float observations. We filtered monthly time series (Fig. 1b) using a running annual mean. Harmonic analyses of zonal velocity and temperature (Figs 1b and 3) were performed by applying a linear regression model in a least-squares sense to the data. We approximated the degrees of freedom used for the calculation of the standard error of the resulting amplitudes as the length of the time series divided by a quarter of the deep-jet oscillation period. The significance of the correlation (Fig. 1a) was obtained using surface wind and rainfall time series of the same length as the microwave optimally interpolated SST with corresponding degrees of freedom. Sources and time intervals of all data sets used in this study are given in Supplementary Table 1.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 November 2010; accepted 21 March 2011.

Published online 18 May 2011.

- Giannini, A., Saravanan, R. & Chang, P. Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. *Science* **302**, 1027–1030 (2003).
- Chang, P., Fang, Y., Saravanan, R., Ji, L. & Seidel, H. The cause of the fragile relationship between the Pacific El Niño and the Atlantic Niño. *Nature* **443**, 324–328 (2006).
- Czaja, A., van der Vaart, P. & Marshall, J. A diagnostic study of the role of remote forcing in tropical Atlantic variability. *J. Clim.* **15**, 3280–3290 (2002).
- Carton, J. A., Cao, X., Giese, B. S. & da Silva, A. M. Decadal and interannual SST variability in the tropical Atlantic Ocean. *J. Phys. Oceanogr.* **26**, 1165–1175 (1996).
- Chang, P., Ji, L. & Li, H. A decadal climate variation in the tropical Atlantic Ocean from thermodynamic air-sea interactions. *Nature* **385**, 516–518 (1997).

- Zebiak, S. E. Air-sea interaction in the equatorial Atlantic region. *J. Clim.* **6**, 1567–1586 (1993).
- Carton, J. A. & Huang, B. Warm events in the tropical Atlantic. *J. Phys. Oceanogr.* **24**, 888–903 (1994).
- Chang, P. *et al.* Climate fluctuations of tropical coupled systems — The role of ocean dynamics. *J. Clim.* **19**, 5122–5174 (2006).
- Kushnir, Y., Robinson, W. A., Chang, P. & Robertson, A. W. The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability. *J. Clim.* **19**, 5949–5970 (2006).
- Johnson, G. C. & Zhang, D. Structure of the Atlantic Ocean equatorial deep jets. *J. Phys. Oceanogr.* **33**, 600–609 (2003).
- Bunge, L., Provost, C., Hua, B. L. & Kartavtseff, A. Variability at intermediate depths at the equator in the Atlantic Ocean in 2000–06: annual cycle, equatorial deep jets, and intraseasonal meridional velocity fluctuations. *J. Phys. Oceanogr.* **38**, 1794–1806 (2008).
- Johnson, G. C., Kunze, E., McTaggart, K. E. & Moore, D. W. Temporal and spatial structure of the equatorial deep jets in the Pacific Ocean. *J. Phys. Oceanogr.* **32**, 3396–3407 (2002).
- Luyten, J. R. & Swallow, J. C. Equatorial undercurrents. *Deep-Sea Res.* **23**, 999–1001 (1976).
- Ruiz-Barradas, A., Carton, J. A. & Nigam, S. Structure of interannual-to-decadal climate variability in the tropical Atlantic sector. *J. Clim.* **13**, 3285–3297 (2000).
- Wang, F. & Chang, P. A linear stability analysis of coupled tropical Atlantic variability. *J. Clim.* **21**, 2421–2436 (2008).
- Ding, H., Keenlyside, N. S. & Latif, M. Equatorial Atlantic interannual variability: the role of heat content. *J. Geophys. Res.* **115**, C09020 (2010).
- Jochum, M., Murtugudde, R., Malanotte-Rizzoli, P. & Busalacchi, A. in *Earth Climate: The Ocean-Atmosphere Interaction* (eds Wang, C., Xie, S.-P. & Carton, J. A.) 181–188 (Geophys. Monogr. Ser. 147, American Geophysical Union, 2004).
- Philander, S. G. H. & Pacanowski, R. C. Response of equatorial oceans to periodic forcing. *J. Geophys. Res.* **86**, 1903–1916 (1981).
- Lebedev, K. V., Yoshinari, H., Maximenko, N. A. & Hacker, P. W. YoMaHa'07: Velocity Data Assessed from Trajectories of Argo Floats at Parking Level and at the Sea Surface. IPRC Technical Note 4 (International Pacific Research Center, 2007).
- Firing, E., Wijffels, S. E. & Hacker, P. Equatorial subthermocline currents across the Pacific. *J. Geophys. Res.* **103**, 21413–21423 (1998).
- Ponte, R. M. & Luyten, J. R. Deep velocity measurements in the western equatorial Indian Ocean. *J. Phys. Oceanogr.* **20**, 44–52 (1990).
- d'Orgeville, M., Hua, B. L. & Sasaki, H. Equatorial deep jets triggered by a large vertical scale variability within the western boundary layer. *J. Mar. Res.* **65**, 1–25 (2007).
- Hua, B. L. *et al.* Destabilization of mixed Rossby gravity waves and the formation of equatorial zonal jets. *J. Fluid Mech.* **610**, 311–341 (2008).
- Eden, C. & Dengler, M. Stacked jets in the deep equatorial Atlantic Ocean. *J. Geophys. Res.* **113**, C04003 (2008).
- McPhaden, M. J., Proehl, J. A. & Rothstein, L. M. The interaction of equatorial Kelvin waves with realistically sheared zonal currents. *J. Phys. Oceanogr.* **16**, 1499–1515 (1986).
- Brown, G. L. & Sutherland, B. R. Internal wave tunnelling through non-uniformly stratified shear flow. *Atmosphere-Ocean* **45**, 47–56 (2007).
- Gouretski, V. V. & Koltermann, K. P. *WOCE Global Hydrographic Climatology*. Report 35 (Bundesamt für Seeschifffahrt und Hydrographie, 2004).
- Cane, M. A. & Moore, D. W. A note on low-frequency equatorial basin modes. *J. Phys. Oceanogr.* **11**, 1578–1584 (1981).
- Keenlyside, N. S. & Latif, M. Understanding equatorial Atlantic interannual variability. *J. Clim.* **20**, 131–142 (2007).
- Okumura, Y. & Xie, S.-P. Some overlooked features of tropical Atlantic climate leading to a new Niño-like phenomenon. *J. Clim.* **19**, 5859–5874 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study was supported by the German Federal Ministry of Education and Research as part of the co-operative project 'North Atlantic' and by the German Science Foundation as part of the Sonderforschungsbereich 754 'Climate-Biogeochemistry Interactions in the Tropical Ocean'. The contribution of J.M.T. was facilitated by support from the Woods Hole Oceanographic Institution's Columbus O'Donnell Iselin Chair for Excellence in Oceanography. We thank J. Fischer for mooring planning and field-work participation, F. Ascani for discussion and S.-H. Didwischus for data processing. This study uses PIRATA velocity and temperature data provided through the TAO project office, Argo float drift data provided by APDRC/IPRC¹⁹, rainfall data from the Global Precipitation Climatology Project, Met Office Hadley Centre and microwave optimally interpolated SST data, NCEP/NCAR reanalysis wind data, and AVISO sea level anomaly data (Supplementary Table 1).

Author Contributions P.B. led the project and designed the study including sea-going work and data analysis. A.F. and V.H. processed and analysed moored velocity, Argo float and satellite data. J.M.T. performed moored profiler measurements, its data processing and its analysis. P.B., M.D. and R.J.G. led the drafting of the manuscript. All authors contributed to the interpretation of the results and provided substantial input to the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.B. (pbrandt@ifm-geomar.de).

METHODS

Surface zonal velocity anomaly at the Equator, averaged between 35° W and 15° W (Fig. 1b), was calculated by applying a second-order fit in latitude to monthly mean meridional sea level anomaly distributions between 1° S and 1° N and evaluating equatorial geostrophy using the obtained curvature. Mean zonal velocities from Argo float drifts between 950 and 1,050 m (Fig. 1b) were derived by removing outliers using a standard-deviation criterion and averaging over time (1-yr period) and space (from 1° S to 1° N and from 35° W to 15° W). The standard error of the nominal 1,000-m zonal velocities (Fig. 1b) was calculated by dividing their standard deviation by the square root of the number of float observations. Monthly time series (Fig. 1b) were filtered using a running annual mean. The dominant period of these time series was estimated by calculating the variance explained by a plane-wave fit (Supplementary Fig. 1).

In the subsurface temperature and velocity time series from PIRATA buoys and subsurface moorings, which are used to produce Fig. 3, data gaps are present. Here monthly time series were derived by monthly averaging and subtracting a mean annual cycle.

Harmonic analyses of zonal velocity and temperature time series (Figs 1b and 3b, c) were performed by fitting the following linear regression model in a least-squares sense to the monthly data:

$$\mathbf{d}_m = g\mathbf{\beta} = \beta_1 \mathbf{I}_N + \beta_2 \cos(\omega \mathbf{t}) + \beta_3 \sin(\omega \mathbf{t})$$

Here \mathbf{t} is the time vector corresponding to the data vector, \mathbf{d} , both of which are of length N ; $\cos(\omega \mathbf{t})$ and $\sin(\omega \mathbf{t})$ are the vectors whose elements are the cosines and sines of the elements of $\omega \mathbf{t}$, respectively; $\omega = 2\pi/p$ is the angular frequency, where p is the period; g is the model matrix; $\mathbf{\beta}$ is a column vector of scalar model factors (β_1 , β_2 and β_3); and \mathbf{I}_N is a vector of length N whose elements all equal 1. The error matrix is given by

$$\Delta\mathbf{\beta} = \sqrt{\frac{(\mathbf{g}^T \mathbf{g})^{-1} (\mathbf{d} - \mathbf{d}_m)^T (\mathbf{d} - \mathbf{d}_m)}{n - k}}$$

where n is the number of degrees of freedom and $k = 2$ is the number of dependent model factors. The standard errors of the elements of $\mathbf{\beta}$ are given by the diagonal elements of $\Delta\mathbf{\beta}$. The degrees of freedom used for the calculation of the standard error of the resulting amplitudes were approximated as the length of the time series divided by a quarter of the deep-jet oscillation period, and are $n = 14$ for ATL3 SST (HadISST), $n = 10$ for ATL3 SST (microwave optimally interpolated SST), $n = 14$

for geostrophic zonal velocity anomaly, $n = 10$ for the Argo float drift data (Fig. 1b and Supplementary Table 2), and $n = 5$ to $n = 7$ for the moored zonal velocity and subsurface temperature data (Fig. 3) varying with depth owing to data gaps. The phase errors (Fig. 3c) are maximum errors derived using the standard errors of the model factors ($\Delta\beta_2$ and $\Delta\beta_3$) by applying linear error propagation for an arbitrary phase lag.

The significance of the correlation (Fig. 1a) is obtained using surface wind and rainfall time series of the same length as the microwave optimally interpolated SST data series (Fig. 1b), which are additionally 270-d low-pass-filtered and have $n = 10$. Sources and periods of all data sets used are given in Supplementary Tab. 1.

Equatorial zonal velocities from 1,000-m Argo float drift data acquired between 1° S and 1° N were plotted as functions of time and longitude in Fig. 4. The data model

$$\mathbf{d}_m = \hat{U} \sin(k\mathbf{x} - \omega \mathbf{t} - \phi \mathbf{I}_N)$$

was applied to the observed zonal velocities. Here \hat{U} is the zonal velocity amplitude, $\sin(k\mathbf{x} - \omega \mathbf{t} - \phi \mathbf{I}_N)$ is the vector whose elements are the sines of the elements of $k\mathbf{x} - \omega \mathbf{t} - \phi \mathbf{I}_N$, \mathbf{x} is the space vector in the zonal direction corresponding to the data vector, k is the zonal wavenumber and ϕ is the phase. By maximizing the variance explained by the fit, propagation characteristics of the dominant interannual variability were obtained (Supplementary Fig. 4). In the Atlantic, this fit explains about 28% of the variance of the equatorial zonal velocity from Argo float drift data after subtracting the annual and semi-annual cycles. In the Pacific, the strongest interannual signal (which explains only 8% of the variance) is found at a period of 740 d with a zonal wavelength of 56×10^3 km. The associated phase velocity corresponds to a first-baroclinic-mode Rossby wave that is very probably forced by the wind (Supplementary Fig. 4). Uncertainties in period and wavelength were estimated by a non-parametric bootstrap procedure where a number of resamples was constructed by random sampling with replacement (Supplementary Fig. 5).

Moored velocity data were acquired using acoustic Doppler current profilers, different single-point current meters and a moored profiler (Figs 2 and 3 and Supplementary Fig. 3). The oceanic variability on short timescales clearly exceeds the measurement accuracy of the different instruments. Owing to a ballasting error, the moored profiler was deployed 'light' and suffered loss of drive-wheel traction over time, resulting in truncation of the down-going profiles as time progressed (Fig. 2).

Immunogenicity of induced pluripotent stem cells

Tongbiao Zhao¹, Zhen-Ning Zhang¹, Zhili Rong¹ & Yang Xu¹

Induced pluripotent stem cells (iPSCs), reprogrammed from somatic cells with defined factors, hold great promise for regenerative medicine as the renewable source of autologous cells^{1–5}. Whereas it has been generally assumed that these autologous cells should be immune-tolerated by the recipient from whom the iPSCs are derived, their immunogenicity has not been vigorously examined. We show here that, whereas embryonic stem cells (ESCs) derived from inbred C57BL/6 (B6) mice can efficiently form teratomas in B6 mice without any evident immune rejection, the allogeneic ESCs from 129/SvJ mice fail to form teratomas in B6 mice due to rapid rejection by recipients. B6 mouse embryonic fibroblasts (MEFs) were reprogrammed into iPSCs by either retroviral approach (ViPSCs) or a novel episomal approach (EiPSCs) that causes no genomic integration. In contrast to B6 ESCs, teratomas formed by B6 ViPSCs were mostly immune-rejected by B6 recipients. In addition, the majority of teratomas formed by B6 EiPSCs were immunogenic in B6 mice with T cell infiltration, and apparent tissue damage and regression were observed in a small fraction of teratomas. Global gene expression analysis of teratomas formed by B6 ESCs and EiPSCs revealed a number of genes frequently over-expressed in teratomas derived from EiPSCs, and several such gene products were shown to contribute directly to the immunogenicity of the B6 EiPSC-derived cells in B6 mice. These findings indicate that, in contrast to derivatives of ESCs, abnormal gene expression in some cells differentiated from iPSCs can induce T-cell-dependent immune response in syngeneic recipients. Therefore, the immunogenicity of therapeutically valuable cells derived from patient-specific iPSCs should be evaluated before any clinic application of these autologous cells into the patients.

To vigorously examine the immunogenicity of cells derived from iPSCs, we took advantage of the capability of ESCs and iPSCs to form teratomas in mice that allows the simultaneous evaluation of the immunogenicity of various cell types derived from them. Whereas B6 ESCs could efficiently form teratomas in B6 mice without any evidence of immune rejection as indicated by the lack of any detectable CD4⁺ T cell infiltration, a hallmark of immune rejection, the allogeneic 129/SvJ (129) ESCs were rapidly rejected before forming detectable teratomas in the same B6 recipients with massive infiltration of CD4⁺ T cells into one detectable teratomas formed by 129 ESCs (Fig. 1a–d, Supplementary Fig. 1). The CD4⁺ cells were not directly differentiated from the implanted ESCs because no CD4⁺ cells were detectable in any examined teratomas formed by B6 and 129 ESCs in severe combined immunodeficient (SCID) mice (Fig. 1d). B6 and 129 ESCs had similar proliferation rates and both could efficiently form teratomas in SCID mice (Supplementary Fig. 1a–e). Therefore, these findings validate the feasibility to use this teratomas formation assay to evaluate the immunogenicity of iPSC derivatives *in vivo*.

We initially established ViPSCs from B6 MEFs with the cocktails of retrovirus expressing either three (Oct4/Sox2/Klf4) or four (Oct4/Sox2/myc/Klf4) reprogramming factors as described¹. The subcloned ViPSCs had normal karyotypes, expressed ESC-specific surface markers and pluripotency genes, and were pluripotent as indicated by their capability to form teratomas in SCID mice and contribute to adult chimaeric mice (Supplementary Fig. 2a–g). Four independent

iPSC clones, two reprogrammed with three factors (V3-1 and V3-3) and two with four factors (V4-1 and V4-2), were selected for further analysis (Supplementary Fig. 2h). Most implanted B6 ViPSCs failed to form detectable teratomas or formed teratomas that were subsequently immune-rejected with T cell infiltration and massive necrosis (Supplementary Fig. 3a–e). The teratomas that did not undergo apparent regression were also infiltrated with CD4⁺ T cells with apparent

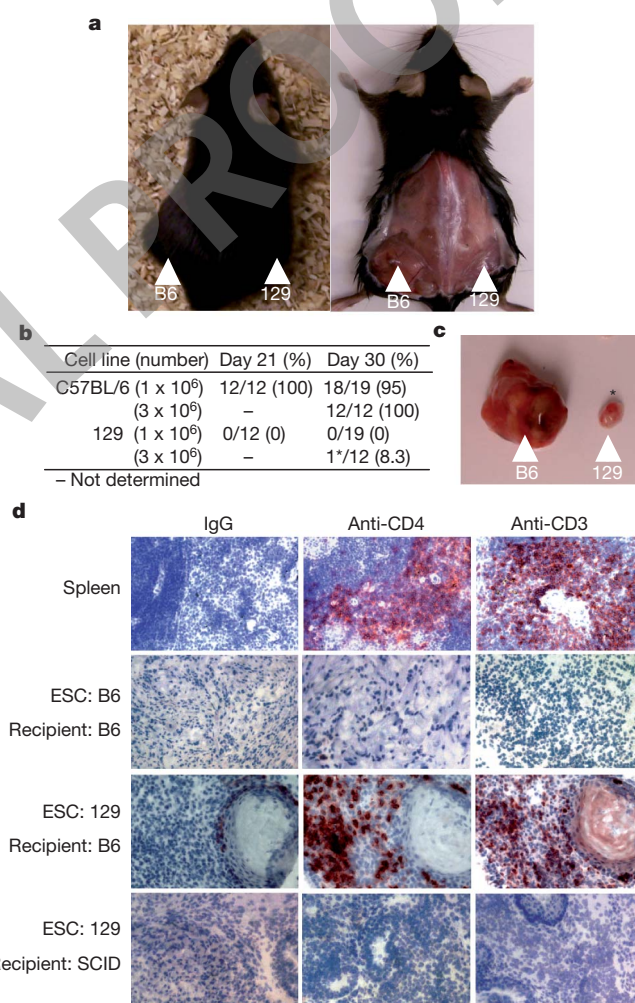


Figure 1 | Immunogenicity of syngeneic and allogeneic ESCs in male B6 mice. **a**, B6 but not 129 ESCs can efficiently form teratomas in B6 mice after subcutaneous injection. The teratomas shown is 30 days after implantation. **b**, Summary of teratomas formation by ESCs in B6 mice 21 and 30 days after implantation. Only one small teratoma formed by 129 ESCs was detected (asterisk) and is shown in **c**. **d**, Infiltration of T cells was detected in the teratomas formed by 129 ESCs but not the ones formed by B6 ESCs in B6 mice. T cells were identified by anti-CD4 and anti-CD3 antibodies. Sections from the spleen and teratomas formed by 129 ESCs in SCID mice were used as positive and negative controls.

¹Section of Molecular Biology, Division of Biological Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0322, USA.

necrosis within parts of the tumour (Supplementary Fig. 3d, e). Therefore, cells derived from B6 ViPSCs are highly immunogenic in B6 mice.

Recent studies have shown the existence of T cells specific for the cells expressing Oct4 in the periphery⁶. Therefore, the reactivation of Oct4 expression in cells differentiated from B6 ViPSCs could induce immune responses in B6 mice (Supplementary Fig. 2i). To address this issue, we developed a novel episomal approach to reprogram B6 MEFs into EiPSCs that express ESC markers and pluripotency genes as well as contribute to adult chimaeric mice (Fig. 2a–e). Extensive Southern blotting analysis demonstrated that some EiPSC clones (1E12, 1E13, 3E1) had lost the episomal vector and harboured no random integration of the reprogramming vector (Fig. 2f). The expression cassette was excised from the genome of 2E2 iPSC clone that harboured one random integration of the episomal vector by transient expression of Cre enzyme (Supplementary Fig. 4).

EiPSCs had normal karyotypes and efficiently formed teratomas in B6 mice. However, the majority of teratomas derived from EiPSCs of both early and late passages showed apparent infiltration of T cells (Figs 3a, d and Supplementary Fig. 5). In addition, apparent tumour regression with extensive tissue necrosis was detected in 10% of teratomas formed by EiPSCs in B6 mice within 2 months of implantation (Fig. 3b, c). No apparent tumour regression was observed in the majority of the teratomas formed by EiPSCs in B6 mice before they

reached the allowed maximal size (Fig. 3c). Therefore, we concluded that cells derived from B6 EiPSCs can be immunogenic in B6 recipients, but their overall immunogenicity is lower than the cells derived from B6 ViPSCs.

To determine the generality of our conclusion, two independently generated integration-free B6 iPSC lines, which were reprogrammed from B6 MEFs with a plasmid vector expressing Oct4/Sox2/Myc/Klf4 (ref. 7), were implanted into B6 mice. T cell infiltration was observed in most teratomas formed by these B6 iPSCs in B6 mice, some of which also exhibit tissue necrosis (Supplementary Fig. 6). In addition, a small fraction of teratomas had undergone apparent regression by 40 days after implantation. These findings support the conclusion that cells derived from iPSCs are immunogenic in syngeneic recipients.

To understand the basis of this immunogenicity, the profile of gene expression in teratomas derived from B6 ESCs and EiPSCs revealed a number of genes overexpressed in teratomas derived from B6 EiPSCs (Supplementary Fig. 7a). Expression analysis of six regressing teratomas formed by two independent B6 EiPSCs in B6 mice indicated that 9 of the 23 tested genes (*Lcel1f*, *Spt1*, *Cyp3a11*, *Zg16*, *Lce3a*, *Chi3L4*, *Olr1*, *Retn*, *Hormad1*) were commonly overexpressed in these teratomas (Fig. 4a). *Hormad1* has been identified as a tumour antigen and *Spt1* as a tissue-specific antigen^{8,9}.

To test the possibility that the abnormal expression of these genes in teratomas derived from B6 iPSCs contributes to their immunogenicity

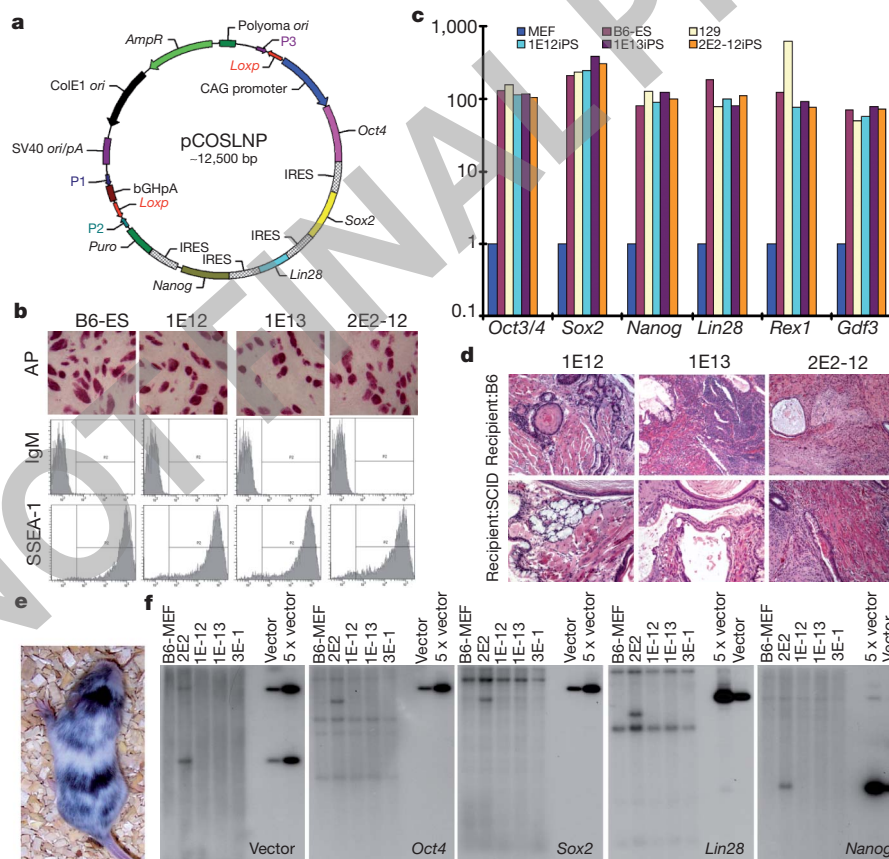
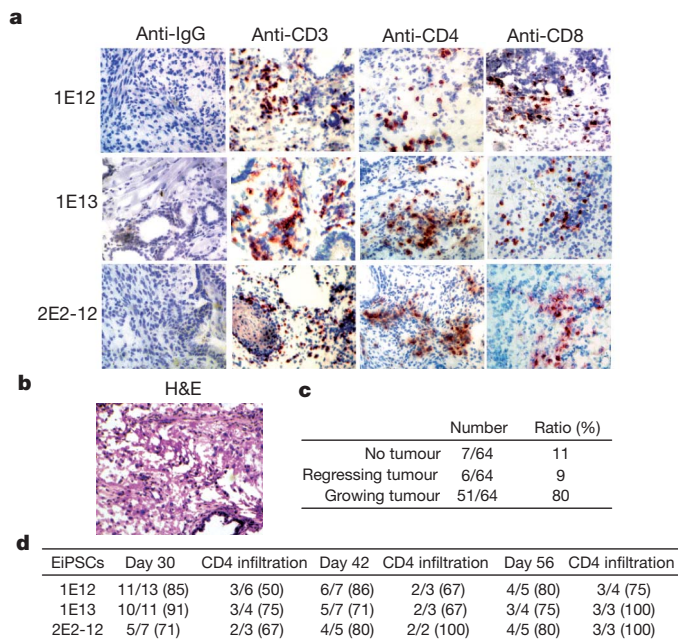


Figure 2 | A new episomal approach to generate EiPSCs from B6 MEFs. **a**, Diagram of the episomal vector that expresses the four reprogramming factors (Oct4/Sox2/Nanog/Lin28) and puromycin resistance gene from one messenger RNA separated by IRES sequences. The entire expression cassette is flanked with LoxP sites. **b**, **c**, EiPSCs were positive for alkaline phosphatase (AP) and SSEA-1 (**b**) and expressed pluripotency genes to the same levels as those of B6 ESCs as determined by quantitative real-time PCR (**c**). The mRNA levels in MEFs are arbitrarily set to 1. **d**, EiPSCs form teratomas in both B6 mice

(top panel) and SCID mice (bottom panel). **e**, EiPSCs can contribute to adult chimaeric mice after injecting into the blastocysts derived from albino mice. **f**, Southern blotting analysis indicates no random integration of the episomal vectors in EiPSC clones 1E-12, 1E-13 and 3E-1. Clone 2E2 has one copy of the episomal vector integrated into the genome. Genomic DNA derived from iPSCs was digested with BamHI and hybridized to various probes that together cover the entire episomal vector.



in B6 mice, seven such genes were ectopically expressed in B6 ESCs and their derived teratomas under the control of the ubiquitously active CAG promoter/enhancer (Supplementary Fig. 7b). Like B6 ESCs, over 90% of implants of B6 ESCs with empty vector as well as

Figure 3 | Cells derived from B6 EiPSCs can be immunogenic in B6 mice. **a**, T-cell infiltration was detected in the majority of teratomas formed by B6 EiPSCs in male B6 mice. 2E2-12 iPSCs is a subclone of 2E2 clone after LoxP/Cre-mediated deletion of the reprogramming factor expression cassette from the integrated copy of episomal vector. **b**, Tissue necrosis was detected in the regressing teratomas formed by B6 EiPSCs in male B6 mice. H&E, haematoxylin and eosin staining. **c**, Summary of teratoma formation by B6 EiPSCs in male B6 mice. **d**, Summary of teratoma formation and CD4⁺ T cell infiltration at different time points after implantation of EiPSCs in male B6 mice.

transgenic *Lce1f*-B6 ESC and *Retn*-B6 ESCs formed teratomas in B6 mice (Fig. 4b). In contrast, over 80% of *Zg16*-B6 ESC implants and 50% of *Hormad1*- or *Cyp3a11*-B6 ESC implants failed to form visible teratomas in B6 mice (Fig. 4b). Extensive T cell infiltration and widespread necrosis were detected in the teratomas formed by *Zg16*- and *Hormad1*-B6 ESCs in B6 mice but rarely detectable in the teratomas derived from *Lce1f*- and *Retn*-B6 ESCs in B6 mice (Fig. 4c, d). To rule out the possibility that the regression of the teratomas formed by *Zg16*- and *Hormad1*-B6 ESCs in B6 mice is secondary to the abnormal proliferation or cell death induced by the ectopic expression of these genes, the proliferation and survival of *Zg16*- and *Hormad1*-B6 ESCs were identical to B6 ESCs (Supplementary Fig. 7c, e). In addition, the weight of the teratomas formed by *Zg16*- and *Hormad1*-B6 ESCs in SCID mice was similar to that of B6 ESCs (Supplementary Fig. 7d).

To identify the immune responses against the cells derived from iPSCs, we used CD4^{-/-} and CD8^{-/-} B6 mice to examine the importance of T cells in the immune rejection. The robust immune rejection of the teratomas formed by B6 ViPSCs as well as *Zg16*- and *Hormad1*-B6 ESCs in B6 mice was abolished in both CD4^{-/-} and CD8^{-/-} B6 mice

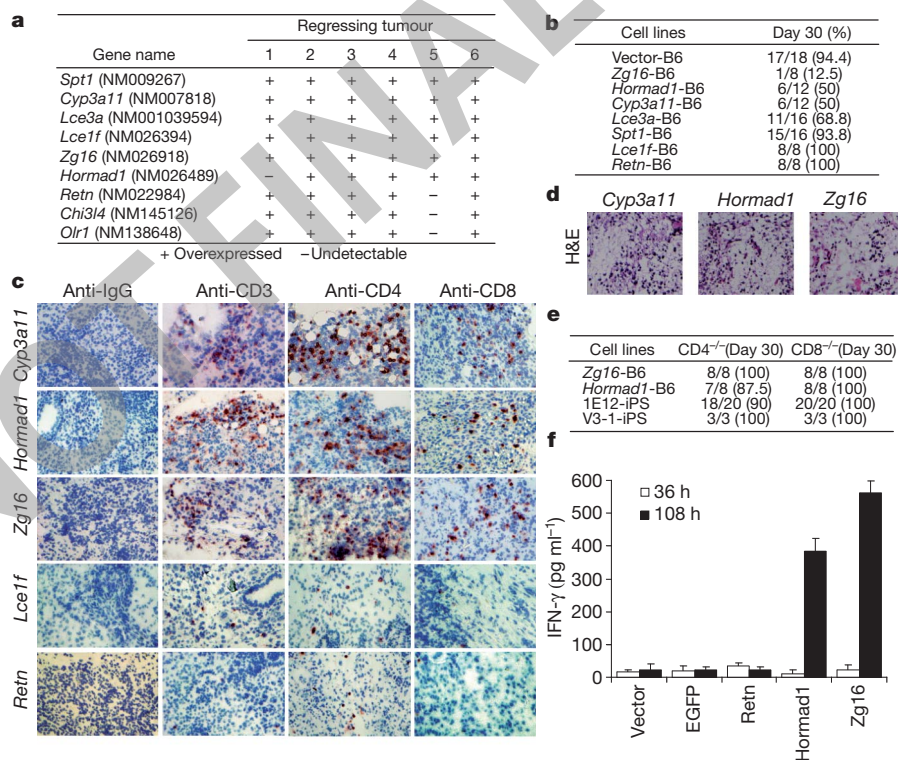


Figure 4 | Abnormal overexpression of some proteins contributes directly to the immunogenicity of cells derived from B6 EiPSC in B6 mice. **a**, Nine genes were found to be commonly overexpressed in six regressing teratomas formed by two B6 EiPSCs. The expression of 23 genes identified as overexpressed in EiPSC-derived teratomas by microarray analysis was analysed by real-time PCR. **b**, Summary of teratoma formation by various transgenic B6 ESCs in male B6 mice. **c**, Extensive infiltration of T cells in the teratomas formed by *Cyp3a11*-, *Hormad1*- and *Zg16*-B6 ESCs in B6 mice. Few infiltrating T cells were detectable in the teratomas formed by *Lce1f*- and *Retn*-B6 ESCs in

B6 mice. Representative images are shown. **d**, Extensive necrosis is present in teratomas formed by *Cyp3a11*-, *Hormad1*- and *Zg16*-B6 ESCs in B6 mice. **e**, The immune rejection of the teratomas formed by *Hormad1*-B6 ESCs, *Zg16*-B6 ESCs, B6 EiPSCs and B6 ViPSCs is abolished in CD4^{-/-} or CD8^{-/-} B6 mice. **f**, IFN- γ release assay to detect the presence of primed T cells specific for cells expressing *Hormad1* and *Zg16* in B6 mice harbouring the teratomas formed by EiPSCs. Each data point represents the mean of duplicate cultures. Consistent data are obtained from three independent experiments.

(Fig. 4e). In addition, no regression of teratomas formed by EiPSCs in CD4^{-/-} or CD8^{-/-} B6 mice was observed. Therefore, both CD4⁺ helper T cells and CD8⁺ cytotoxic T cells are critical for this immune rejection. These findings also indicate that the innate immunity does not have an important role in the immune rejection of the cells derived from iPSCs.

To further determine whether the abnormal expression of Hormad1 and Zg16 in teratomas formed by EiPSCs directly activates T-cell responses in B6 mice, we performed the IFN- γ releasing assay that measures the antigen-specific activation of *in-vivo*-primed T cells¹⁰. Dendritic cells purified from B6 mice were transfected with either empty expression vector or vectors expressing Zg16, Hormad1, Retn or EGFP (enhanced green fluorescent protein). LPS-matured dendritic cells expressing Hormad1 or Zg16 but not the dendritic cells expressing Retn or EGFP could induce IFN- γ production from purified T cells, indicating the presence of primed T cells specific for cells expressing Hormad1 or Zg16 in B6 mice harbouring the teratomas formed by EiPSCs (Fig. 4f). Although these findings did not identify the specific peptides responsible for activating T cells, they demonstrate that the abnormal expression of Hormad1 and Zg16 contributes directly to the immunogenicity of the cells derived from EiPSCs in syngeneic recipients. Hormad1 was also overexpressed in most teratomas formed by four independently generated integration-free iPSCs reprogrammed with adenoviral vectors, recombinant proteins or plasmid vectors^{7,11,12}. In addition, Zg16 was overexpressed in most teratomas formed by iPSCs reprogrammed with recombinant proteins. Therefore, the abnormal expression of such immunogenic proteins could represent a common mechanism to induce T cell-mediated immune responses to cells derived from iPSCs.

Our findings indicate that some cells derived from iPSCs can be immunogenic in syngeneic recipients. The T-dependent immune response is likely due to the abnormal expression of antigens not expressed during normal development or differentiation of ESCs, leading to the break of peripheral tolerance. The expression of these minor antigens could be due to the subtle yet apparent epigenetic difference between iPSCs and ESCs^{13–20}. In addition, recently discovered mutations in the coding sequences of iPSCs could also contribute to the immunogenicity of iPSC derivatives²¹. Therefore, for the clinic development of iPSCs, current reprogramming technology needs to be optimized to minimize the epigenetic difference between iPSCs and ESCs. The *in vivo* immunogenicity test described here can provide a robust screening platform for improving the reprogramming technology.

METHODS SUMMARY

Mice. B6 mice and ESCs were purchased from The Jackson Laboratory. Only male mice were used in the transplantation studies. All animal experiments were performed in accordance with relevant guidelines and regulations, and approved by the Institutional Animal Care and Use Committee (IACUC).

iPSC generation and characterization. MEFs were isolated from B6 embryo as described²². For ViPSC production, MEFs were transduced with retrovirus cocktail as described¹. For EiPSCs generation, MEFs were transfected with the episomal vector expressing the reprogramming factors. The lack of random integration of the episomal vector was confirmed by Southern blotting analysis with a combination of probes that cover the entire episomal vector.

Interferon- γ releasing assay. Dendritic cells from B6 mice were isolated and transfected with expression vectors. The transfected dendritic cells were matured by lipopolysaccharide (LPS) treatment for 12 h. T cells were purified from the pooled spleens and lymph nodes of five B6 mice harbouring the teratomas formed by EiPSCs and co-cultured with LPS-matured dendritic cells. Supernatant were collected at indicated time point to determine the IFN- γ levels.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 July 2010; accepted 19 April 2011.

Published online 13 May 2011.

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
2. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
3. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
4. Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
5. Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **105**, 2883–2888 (2008).
6. Dhodapkar, K. M. *et al.* Natural immunity to pluripotency antigen OCT4 in humans. *Proc. Natl Acad. Sci. USA* **107**, 8718–8723 (2010).
7. Jincho, Y. *et al.* Generation of genome integration-free induced pluripotent stem cells from fibroblasts of C57BL/6 mice without c-Myc transduction. *J. Biol. Chem.* **285**, 26384–26389 (2010).
8. Chen, Y. T. *et al.* Identification of CT46/HORMAD1, an immunogenic cancer/testis antigen encoding a putative meiosis-related protein. *Cancer Immun.* **5**, 9 (2005).
9. Kont, V. *et al.* Modulation of Aire regulates the expression of tissue-restricted antigens. *Mol. Immunol.* **45**, 25–33 (2008).
10. Guernonprez, P., Valladeau, J., Zitvogel, L., Théry, C. & Amigorena, S. Antigen presentation and T cell stimulation by dendritic cells. *Annu. Rev. Immunol.* **20**, 621–667 (2002).
11. Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. & Hochedlinger, K. Induced pluripotent stem cells generated without viral integration. *Science* **322**, 945–949 (2008).
12. Zhou, H. *et al.* Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* **4**, 381–384 (2009).
13. Chin, M. H. *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* **5**, 111–123 (2009).
14. Liu, L. *et al.* Activation of the imprinted *Dlk1-Dio3* region correlates with pluripotency levels of mouse stem cells. *J. Biol. Chem.* **285**, 19483–19490 (2010).
15. Stadtfeld, M. *et al.* Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465**, 175–181 (2010).
16. Doi, A. *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genet.* **41**, 1350–1353 (2009).
17. Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285–290 (2010).
18. Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotechnol.* **28**, 848–855 (2010).
19. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
20. Zhao, T. & Xu, Y. p53 and stem cells: new developments and new concerns. *Trends Cell Biol.* **20**, 170–175 (2010).
21. Gore, A. *et al.* Somatic coding mutations in human induced pluripotent stem cells. *Nature* **471**, 63–67 (2011).
22. Song, H., Hollstein, M. & Xu, Y. p53 gain-of-function cancer mutants induce genetic instability by inactivating ATM. *Nature Cell Biol.* **9**, 573–580 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Abe, S. Ding and K. Hochedlinger for their generous supply of integration-free mouse iPSCs. We thank N. Shastri for his advice on how to identify antigen-specific T cells. We thank J. Fink and Blue Lake of Xu lab as well as UCSD Cancer Center pathologic core for technical support. This work was supported by a NIH grant and an Early Translational Award from California Institute for Regenerative Medicine to Y.X. (ET-01277).

Author Contributions T.Z. and Y.X. designed the experiments, analysed the data and wrote the manuscript. T.Z., Z.-N.Z. and Z.R. executed the experiments under the overall coordination of T.Z.

Author Information The microarray data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE28573. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.X. (yangxu@ucsd.edu).

METHODS

Mice, cell culture and reprogramming episomal vector construction. C57BL/6 (B6) inbred mouse strain and ESCs were purchased from The Jackson Laboratory. Only male mice were used in the transplantation studies of ESCs and iPSCs. All animal experiments were performed in accordance with relevant guidelines and regulations, and approved by the Institutional Animal Care and Use Committee (IACUC). The ESCs and iPSCs were grown on the feeder layer derived from B6 MEFs under standard conditions. The full-length cDNA of *Oct4*, *Sox2*, *Lin28* and *Nanog* was sequentially inserted downstream of the CAG promoter in the episomal vector, separated by IRES sequences (Fig. 2a). The fifth gene, the puromycin resistance gene, is at the 3' end of this mRNA transcript, separated from *Nanog* cDNA by the IRES sequence. This episomal vector is denoted pCOSLNP (CAG–*Oct4*–*Sox2*–*Lin28*–*Nanog*–*Puro*). Two LoxP sites in the same orientation were inserted into the episomal vector flanking the entire expression cassette.

iPSC generation and characterization. MEFs were isolated from B6 embryo as previously described²². For ViPSC production, MEFs were transduced with retrovirus cocktail expressing *Oct4*, *Sox2*, *Klf4* with or without *c-Myc*. The iPSC colonies were picked 18 days after infection as described¹. For EiPSCs generation, MEFs were transfected with pCOSLNP vector using Basic Nucleofector Kit for Primary Mammalian Fibroblasts (Lonza) followed by puromycin selection for 3 days, and then plated on irradiated B6 MEF feeders. Three weeks later, the culture was replated on fresh feeder cells. iPSC colonies were picked 10 to 30 days after replating, and the lack of random integration of the episomal vector was confirmed by Southern blotting analysis with a combination of probes that cover the entire episomal vector.

Quantitative real-time PCR analysis. Total RNA was purified from fibroblasts, ES cells, iPS cells and teratomas with a RNeasy total RNA isolation kit (Qiagen). Total RNA (1 µg) was reversely transcribed into cDNA, which was analysed by quantitative real time PCR analysis as previously described²². The primers used were as follows: *Oct4*F, 5'-GGCTCTCCCATGCATTCAA-3'; *Oct4*R, 5'-TTTAAACCCCAAAGCTCCAGG-3'; *Sox2*F, 5'-AAATCTCCGACGCGAAACG-3'; *Sox2*R, 5'-CCCCAAAAGAAGTCCCAAGA-3'; *Lin28*F, 5'-CTGCTGTAGC GTGATGGTTGA-3'; *Lin28*R, 5'-CCACCCAATGTGTTCTATTGCA-3'; *Nanog*F, 5'-TCGCCATCACACTGACATGA-3'; *Nanog*R, 5'-TGTGCAGAGCA TCTCAGTAGCA-3'; *Rex1*F, 5'-ACGAGTGGCAGTTTCTTCTTGGA-3'; *Rex1*R, 5'-TATGACTCACTCCAGGGGGGCACT-3'; *Gdf3*F, 5'-GATTGCTT TTTCTGCGGTCTGT-3'; *Gdf3*R, 5'-CCAAGTCTTCAGTCGGTTGCT-3'. Primers used for detection of reprogramming factor deletion were as follows: *Oct4*F (43–63), 5'-CCTTCCTTCCCCATGGCGGA-3'; *IRESR1* (53–31), 5'-TTATTCCAAGCGGCTTCGGCCAG-3'; *Sox2*F (1292–1310), 5'-CCCCAG CAGACTTCACATGT-3'; *IRES-R* (221–202), 5'-AGGAAGTCTTCTTCA CGA-3'; *IRESF2* (476–498), 5'-TCGGTGCACATGCTTTACATGTG-3'; *Lin28*R (369–352), 5'-CCGGAACCCCTCCATGTG-3'; *Nanog*RTGA (1131–1111), 5'-TCACACGTCTTCAGGTGCA-3'; *P1*, 5'-CGCCATCTTCTGAAG CTGAATC-3'; *P2*, 5'-ACCGAAAGGAGCGCAGACCCCAT-3'; *P3*, 5'-CCTA CTCAGACAATGCGATGCA-3'; *GAPDH*F, 5'-CCAGTATGACTCCACTCA CG-3'; *GAPDH*R, 5'-GACTCCACGACATACTCAGC-3'; *Lce1f*F, 5'-CTGTA GCCTGGGTCTTG-3'; *Lce1f*R, 5'-GACGATGGCGACGAAGAG-3'; *Spt1R*,

5'-TGAAACTCAGGCAGATAG-3'; *Spt1R*, 5'-TGTCAACGCCACTGTTCT-3'; *Olrl*F, 5'-TGGTGGTCCCTGCTGCTA-3'; *Olrl*R, 5'-ATCCTGCTGAGTAAG GTTCG-3'; *Zg16*F, 5'-CATCACCGCCTTCCGTAT-3'; *Zg16*R, 5'-CGTTGAAA CTTGTGCTGA-3'; *Retn*F, 5'-TCCTGTCCCTGAAGTGC-3'; *Retn*R, 5'-ACG AATGTCCCACGAGCC-3'; *Hormad1*F, 5'-CCAGATTACCAACCACCAG-3'; *Hormad1*R, 5'-TGAAAAGGTGTTGGGACT-3'; *Lce3a*F, 5'-GGCAGTGGTCA GCAGTCT-3'; *Lce3a*R, 5'-TTGGGAAATCCATTAGAAGA-3'; *Cyp3a1*F, 5'-ATCCCATGCTAATAGAC-3'; *Cyp3a1*R, 5'-ATCATCACTGTTGACCCT-3'; *Chi3l4*F, 5'-ATGGCTACACTGGAGAAA-3'; *Chi3l4*R, 5'-TGCTGGAATCCC ACAAT-3'.

Southern blotting analysis. Genomic DNA (10 µg) was digested with BamHI, separated on 1% agarose gel and transferred to a nylon membrane. For the analysis of ViPSCs, the membrane was hybridized to the *Oct4* cDNA probe. For the analysis of EiPSCs, the membrane was hybridized to the cDNA probe of *Oct4*, *Sox2*, *Lin28* and *Nanog* as well as the vector backbone probe.

Teratoma formation and immunohistochemistry analysis. ESCs or iPSCs were collected, washed twice with PBS, and injected subcutaneously into the hind leg region of B6 or SCID mice. One or three million cells were used for each injection. Tumours were measured and surgically removed from the euthanized mice at the indicated time point. Teratomas were fixed either with 4% formaldehyde or frozen in optimal cutting temperature (OCT) compound. Sections were stained with haematoxylin and eosin or with various antibodies such as IgG, anti-CD4, anti-CD3 (BD Biosciences) as we described previously²³.

Microarray assay. Total RNA was purified from the teratomas with an RNeasy total RNA isolation kit (Qiagen). Microarray assay was performed by SeqWright using an Affymetrix Mouse 430A 2.0 chip.

Flow cytometric analysis. About 5×10^5 ESCs or iPSCs were stained for the expression of ESC-specific surface marker with anti-SSEA-1 antibody (Stemgents). Isotype-matched normal antibodies were used as negative controls. The stained cells were analysed by a BD LSR-II using FACSDiva software (Becton Dickinson) as we previously described²².

Interferon-γ releasing assay. To obtain dendritic cells from B6 mice with the dendritic cell purification kit (Miltenyi Biotec), bone marrow cells were isolated from B6 mice and grown in Petri dish at a density of 10^6 cells ml^{-1} in complete medium supplemented with 10 ng ml^{-1} granulocyte/macrophage colony-stimulating factor (GM-CSF) and 5 ng ml^{-1} IL-4 according to the manufacturer's recommendation. On day 9, dendritic cells were purified and transfected with expression vectors using a mouse dendritic cell Nucleofector kit according to the manufacturer's instruction (Lonza). The transfected dendritic cells were matured by LPS treatment for 12 h. T cells were purified from the pooled spleens and lymph nodes of five B6 mice harbouring the teratomas formed by EiPSCs through negative selection using a pan T cell isolation kit (Miltenyi Biotec). Purified T cells (10^6) were immediately co-cultured with LPS-matured dendritic cells (2×10^5). Supernatant were collected at indicated time point to determine the IFN-γ levels using an ELISA kit (Thermo Scientific).

23. Song, H., Chung, S.-K. & Xu, Y. Modeling Disease in Human ESCs Using an Efficient BAC-Based Homologous Recombination System. *Cell Stem Cell* **6**, 80–89 (2010).

Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation

Guillaume Lebon¹, Tony Warne¹, Patricia C. Edwards¹, Kirstie Bennett², Christopher J. Langmead², Andrew G. W. Leslie¹ & Christopher G. Tate¹

Adenosine receptors and β -adrenoceptors are G-protein-coupled receptors (GPCRs) that activate intracellular G proteins on binding the agonists adenosine¹ or noradrenaline², respectively. GPCRs have similar structures consisting of seven transmembrane helices that contain well-conserved sequence motifs, indicating that they are probably activated by a common mechanism^{3,4}. Recent structures of β -adrenoceptors highlight residues in transmembrane region 5 that initially bind specifically to agonists rather than to antagonists, indicating that these residues have an important role in agonist-induced activation of receptors^{5–7}. Here we present two crystal structures of the thermostabilized human adenosine A_{2A} receptor (A_{2A}R-GL31) bound to its endogenous agonist adenosine and the synthetic agonist NECA. The structures represent an intermediate conformation between the inactive and active states, because they share all the features of GPCRs that are thought to be in a fully activated state, except that the cytoplasmic end of transmembrane helix 6 partially occludes the G-protein-binding site. The adenine substituent of the agonists binds in a similar fashion to the chemically related region of the inverse agonist ZM241385 (ref. 8). Both agonists contain a ribose group, not found in ZM241385, which extends deep into the ligand-binding pocket where it makes polar interactions with conserved residues in H7 (Ser 277^{7,42} and His 278^{7,43}; superscripts refer to Ballesteros–Weinstein numbering⁹) and non-polar interactions with residues in H3. In contrast, the inverse agonist ZM241385 does not interact with any of these residues and comparison with the agonist-bound structures indicates that ZM241385 sterically prevents the conformational change in H5 and therefore it acts as an inverse agonist. Comparison of the agonist-bound structures of A_{2A}R with the agonist-bound structures of β -adrenoceptors indicates that the contraction of the ligand-binding pocket caused by the inward motion of helices 3, 5 and 7 may be a common feature in the activation of all GPCRs.

In the simplest model for the conformational dynamics of GPCRs¹⁰ there is an equilibrium between two states, R and R*. The inactive state R preferentially binds inverse agonists and the activated state R* preferentially binds agonists¹¹. Only R* can couple and activate G proteins. Although there are far more complex schemes¹² describing intermediates between R and R*, studies on rhodopsin have indicated that there is only one major conformational change that significantly alters the structure of the receptor³. Thus the structures of dark-state rhodopsin^{13,14} and of opsin^{15,16} are considered to be representative structures for the R and R* state, respectively. Structures of six different GPCRs^{8,13,17–21} in conformations closely approximating the R state have now been determined and it is clear that they are similar to each other, with root mean squared deviation (r.m.s.d.) between any pair of structures in the transmembrane domains being less than 3 Å. As observed in light activation of rhodopsin, the major structural difference between R and R* is the movement of the cytoplasmic ends of helices 5 and 6 away from the receptor core by 5–6 Å, opening up a cleft in the centre of the helix bundle where the carboxy terminus of a G protein can bind¹⁶. Recently,

the structure of an agonist-bound β -adrenoceptor (β_2 -AR) was determined in complex with an antibody fragment (nanobody Nb80)⁵. This structure of β_2 -AR is very similar to the structure of opsin, which indicates that the nanobody mimicked the action of a G protein by maintaining the receptor structure in an activated state. Given the structural similarities between opsin and the β_2 -AR–Nb80 complex, it is likely that the structures of the R* states of other GPCRs are also highly similar. This is consistent with the same heterotrimeric G proteins being able to couple to multiple different receptors²². However, it is unclear whether conserved structures of R and R* indicate that all agonists activate the receptors in an identical fashion. The recent structures of a thermostabilized β_1 -AR bound to four different agonists indicated that a defining feature of agonist binding to this receptor is the formation of a hydrogen bond with Ser^{5,46} on transmembrane helix 5 that accompanies the contraction of the ligand-binding pocket⁷. Here we describe two structures of the adenosine A_{2A} receptor (A_{2A}R) bound to two different agonists, which indicates that the initial action of agonist binding to A_{2A}R has both similarities and differences compared to agonist binding in β -ARs.

The native human A_{2A}R when bound to its endogenous agonist adenosine or to the high-affinity synthetic agonist NECA is unstable in detergent, so crystallization and structure determination relied on using a thermostabilized construct (A_{2A}R-GL31) that contained four point mutations, which markedly improved its thermostability. Pharmacological analysis showed that the mutant receptor bound the five antagonists tested with greatly reduced affinity (1.8–4.3 log units), whereas four agonists bound with similar affinity to the wild-type receptor (Supplementary Fig. 1). However, A_{2A}R-GL31 is only weakly activated by the agonist CGS21680 (Supplementary Fig. 2), which indicates that the thermostabilizing mutations might also decouple high-affinity agonist binding from the formation of R*. The conformation of A_{2A}R-GL31 is not consistent with it being in the fully activated G-protein-coupled state, because we do not observe a 42-fold increase in affinity for NECA binding measured for G_{αs}-coupled A_{2A}R (ref. 23). These data all indicate that A_{2A}R-GL31 is in an intermediate conformation between R and R*, which is consistent with the structural analysis presented later.

The two structures we have determined are of A_{2A}R-GL31 bound to adenosine and NECA with resolutions of 3.0 Å and 2.6 Å, respectively (Supplementary Table 1). Global alignments of the A_{2A}R-GL31 structures with A_{2A}-T4L (A_{2A}R with T4 lysozyme inserted into inner loop 3) bound to the inverse agonist ZM241385 were performed based on those residues in the region of the ligand-binding pocket that show the closest structural homology (Fig. 1 and Supplementary Text). This gave an r.m.s.d. in C α positions of 0.66 Å for the 96 atoms selected, which include all residues involved in binding either adenosine or NECA, with the exception of those in H3. Using this transformation, the adenine-like moiety of the two ligands superimposes almost exactly (r.m.s.d. 0.56 Å). The most significant differences between the two structures are seen in a distortion and a 2 Å shift primarily along the helical axis of H3, a bulge in H5 (resulting from non-helical backbone conformation angles of residues Cys 185 and Val 186) that

¹MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. ²Heptares Therapeutics, BioPark, Broadwater Road, Welwyn Garden City AL7 3AX, UK.

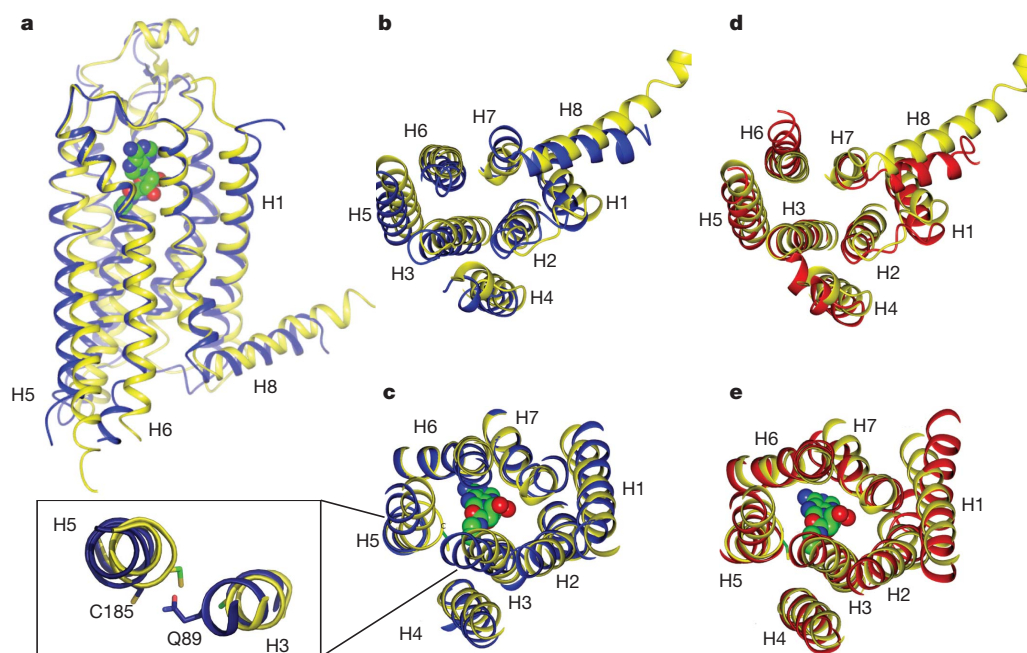


Figure 1 | Structure of the adenosine A_{2A} receptor bound to NECA compared to other GPCR structures. **a**, The structure of NECA-bound A_{2A} R is shown as a cartoon (GPCR) aligned with the structure of A_{2A} -T4L bound to the inverse agonist ZM241385 (blue; PDB code 3EML⁸). NECA is shown as a space-filling model (C, green; N, blue; O, red). **b**, **c**, Sections through the aligned receptors in **a** that highlight the differences in the intracellular face of the

receptors (**b**) and in the ligand-binding pocket (**c**), with the bulge in H5 shown as an inset. **d**, **e**, Alignment of NECA-bound A_{2A} R (yellow) with agonist-bound β_2 -AR-Nb80 (red; PDB code 3P0G⁵) showing the intracellular face of the receptors (**d**) and the ligand-binding pocket (**e**). NECA is shown as a space-filling model in **c** and **e**. The figures were generated using CCP4mg³¹. Analogous alignments to opsin are depicted in Supplementary Fig. 7.

shifts residues into the binding pocket by up to 2 Å and also a change in conformation of the cytoplasmic ends of H5, H6 and H7 (Fig. 1). Comparison of the A_{2A} R-GL31 structure with the agonist-bound β_2 -AR-Nb80 complex indicates that these differences are similar to the conformational changes in the β_2 -AR that are proposed to be responsible for the formation of the R^* state⁵. However, it is unlikely that the structure of A_{2A} R-GL31 represents the fully activated state, because comparison with opsin bound to the C-terminal peptide of the G protein transducin shows that there is insufficient space in A_{2A} R-GL31 for the C terminus of the G protein to bind (Supplementary Fig. 3). This is on the basis of the assumption that all G proteins bind and activate GPCRs in a similar fashion, but given the highly conserved structures of both G proteins and GPCRs this seems a reasonable hypothesis.

The fact that the structure of A_{2A} R-GL31 represents an agonist-binding state is consistent with how A_{2A} R-GL31 was engineered. Thermostabilizing mutations were selected by heating the NECA-bound detergent-solubilised receptor, so the mutations are anticipated to stabilize the agonist-bound state either by stabilizing helix–helix interactions and/or biasing the conformational equilibrium between the agonist-bound R^* state and the agonist-bound R state^{24–26}. The two most thermostabilizing mutations, L48A and Q89A, are in regions of the receptor that are involved in transitions between R and R^* , providing a possible explanation for their thermostabilizing effect (Supplementary Fig. 4). The other two mutations, A54L and T65A, are at the receptor–lipid interface and the reason for their thermostabilizing effect is unclear. Although the overall shape of the ligand-binding pockets of A_{2A} R and β_2 AR are different, the structural similarities with the β_2 -AR-Nb80 (ref. 5) and the structural differences to ZM241385-bound A_{2A} -T4L⁸ indicate that the structure of the binding pocket in A_{2A} R-GL31 is a good representation of the agonist-bound binding pocket of the wild-type receptor (Fig. 1).

Adenosine and NECA bind to A_{2A} R-GL31 in a virtually identical fashion; in addition, the adenine ring in the agonists interacts with A_{2A} R in a similar way to the chemically related triazolo-triazine ring of the inverse agonist ZM241385 (Fig. 2). Thus the hydrogen bonds

between exocyclic adenosine N6 (Supplementary Fig. 5) with both Glu 169 in extracellular loop 2 (EL2) and Asn 253^{6,55} in H6 are similar, with the significant π -stacking interaction with Phe 168 in EL2 also conserved. One of the major structural differences between ZM241385 and the agonists is the presence of a furan substituent on C20 of triazolo-triazine in the inverse agonist, whereas agonists contain a ribose substituent linked to N9 of adenine (Fig. 2 and Supplementary Fig. 5). In ZM241385, the furan group forms a hydrogen bond with Asn 253^{6,55} in H6 and van der Waals contacts with other residues in H3, H5 and H6 (ref. 8). In contrast, the ribose moiety in agonists forms hydrogen bonds with Ser 277^{7,42} and His 278^{7,43} in H7, in addition to van der Waals interactions with other residues in H3 and H6 (Fig. 2). In particular, Val 84^{3,32} has to shift its position upon agonist binding owing to a steric clash with the ribose ring, which may contribute to the 2 Å shift observed in H3 (Fig. 3). These differences in binding between ZM241385 and either adenosine or NECA indicate that the residues that bind uniquely to agonists (Ser 277^{7,42} and His 278^{7,43}) have a key role in the activation of the receptor, as previously shown by mutagenesis studies^{27,28}. This is analogous to the situation in the activation of β_1 -AR, where only full agonists cause the rotamer conformation changes of Ser^{5,46} in H5, whereas the inverse agonist ICI118551 prevents receptor activation by sterically blocking the rotamer change^{7,29}. However, the details of the activation differ in that the critical residues that bind agonists and not antagonists are in H5 in the β_1 -AR, but in H7 in the A_{2A} R (Fig. 4).

Adenosine and NECA activate the A_{2A} R through interactions with H3 and H7 that are absent in the interactions between the receptor and the inverse agonist ZM241385 (Fig. 2). The inward shift of H7, the movement of H3 and the consequent formation of a bulge in H5 are all observed in the structures of agonist-bound A_{2A} R-GL31 and β_2 -AR-Nb80 (Fig. 1). The formation of the bulge in H5 of the β_2 -AR-Nb80 structure was linked to a series of conformational changes that generate the 60° rotation of H6 about Phe 282^{6,44}, resulting in the cytoplasmic end of H6 moving out from the receptor centre and opening the cleft where the C terminus of a G protein is predicted to bind as observed in opsin^{5,6}. There are analogous side-chain movements in A_{2A} R-GL31

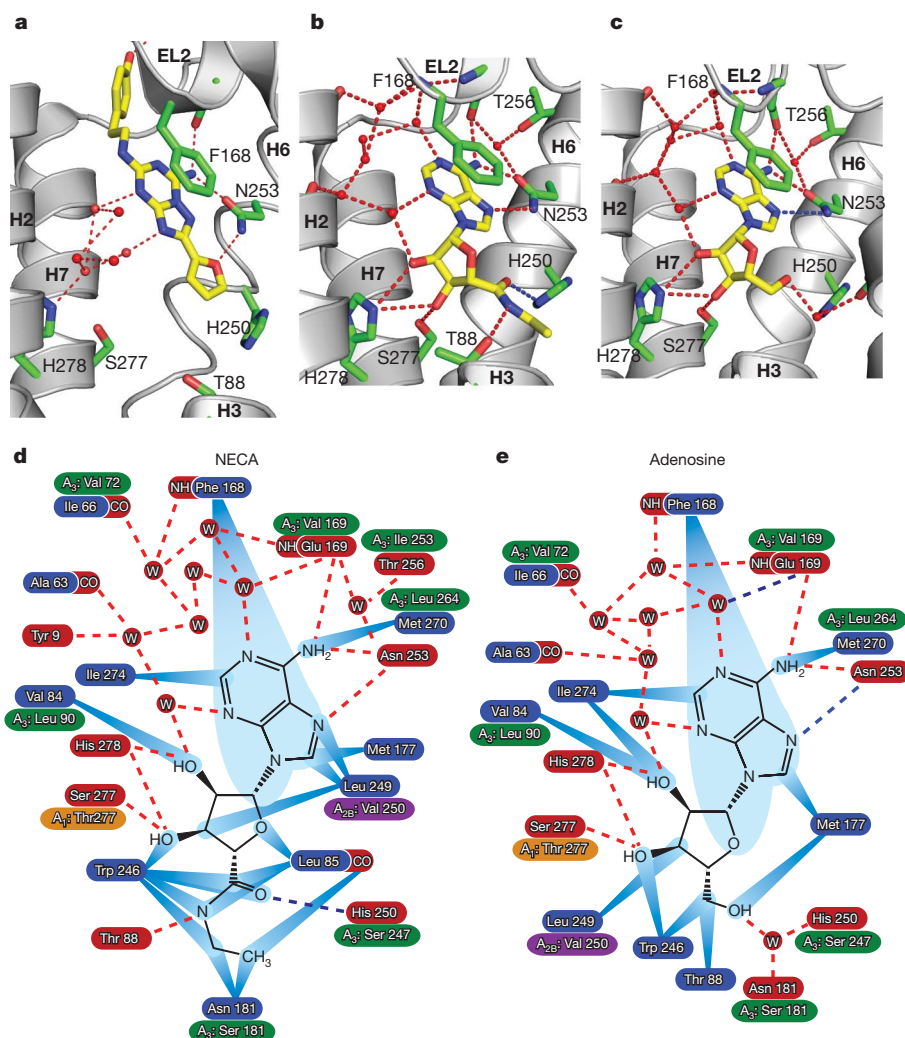


Figure 2 | Comparison of receptor-ligand interactions for A_{2A} R bound to the inverse agonist ZM241385 and the agonists NECA and adenosine. a–c, Structures of the human A_{2A} R in cartoon representation are shown bound to the following ligands: a, ZM241385 (PDB code 3EML⁸); b, NECA; and c, adenosine. d, e, Polar and non-polar interactions involved in agonist binding to A_{2A} R are shown for NECA (d) and adenosine (e). Amino acid residues within 3.9 Å of the ligands are depicted, with residues highlighted in blue making van der Waals contacts (blue rays) and residues highlighted in red making potential hydrogen bonds with favourable geometry (red dashed lines, as identified by HBPLUS, see Methods) or hydrogen bonds with unfavourable geometry (blue dashed lines, donor acceptor distance more than 3.6 Å). Where the amino acid residue differs between the human A_{2A} R and the human A_{1} R, A_{2B} R and A_{3} R, the equivalent residue is shown highlighted in orange, purple or green, respectively. Panels a–c were generated using PyMOL (<http://www.pymol.org/>). Omit densities for the ligands are shown in Supplementary Fig. 6 and densities for water molecules in Supplementary Fig. 8.

that result in a 40° rotation of H6, but the cytoplasmic end of H6 remains partially occluding the G-protein-binding cleft (Supplementary Fig. 3), perhaps because the fully active conformation requires

the binding of G proteins to stabilize it. Interestingly, the structure of β_2 -AR⁶ with a covalently bound agonist is also not in the fully activated R^* conformation, which is only seen after the nanobody Nb80 is bound⁵. The importance of the bulge in H5 in the activation of A_{2A} R is highlighted by how inverse agonists bind. Formation of the H5 bulge results in the inward movement of Cys 185^{5,46} (C β moves by 4 Å), which in turn causes the movement of Val 186 and ultimately a shift of His 250^{6,52} by 2 Å into the ligand-binding pocket, thereby sterically blocking the binding of ZM241385 (Supplementary Fig. 4). Hence, when the inverse agonist binds, it is anticipated that the H5 bulge is unlikely to form owing to the opposite series of events and hence the formation of the R^* state is inhibited.

Thus, in both β -ARs and A_{2A} R, the formation of the H5 bulge seems to be a common action of agonists, whereas inverse agonists seem to prevent its formation. However, the energetic contributions to its formation may be different between the two receptors. In β -ARs there is a major contribution from direct interaction between the agonist and Ser^{5,46}, whereas in A_{2A} Rs, the major interaction seems to come from

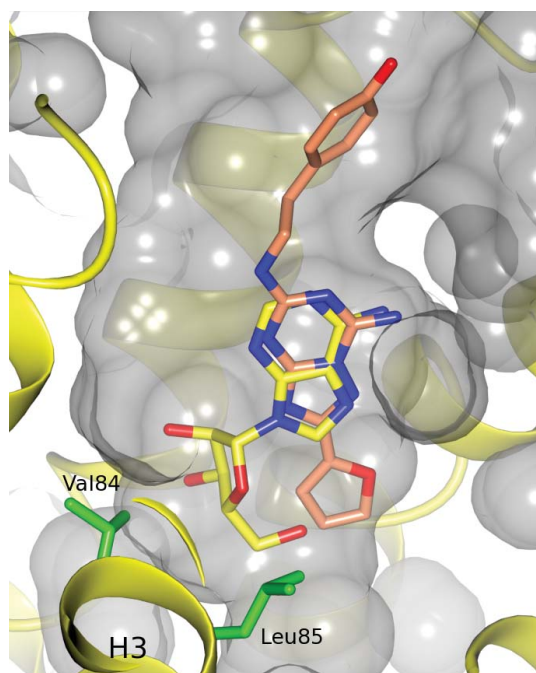


Figure 3 | Positions of adenosine and ZM241385 in the A_{2A} R ligand-binding pocket. The structures of adenosine-bound A_{2A} R-GL31 and ZM241385-bound A_{2A} -T4L were aligned using only atoms from the protein to allow the ligand positions to be compared, with adenosine in yellow and ZM241385 in pink (N, blue; O, red). The ligands are shown in the context of the binding pocket of A_{2A} R-GL31, with transmembrane helices of A_{2A} R-GL31 shown in yellow and the surfaces of the receptor, including the cavity of the ligand binding pocket, shown in grey. The side chains of Val 84 and Leu 85 that interact with the ribose moiety of the agonist are shown in green.

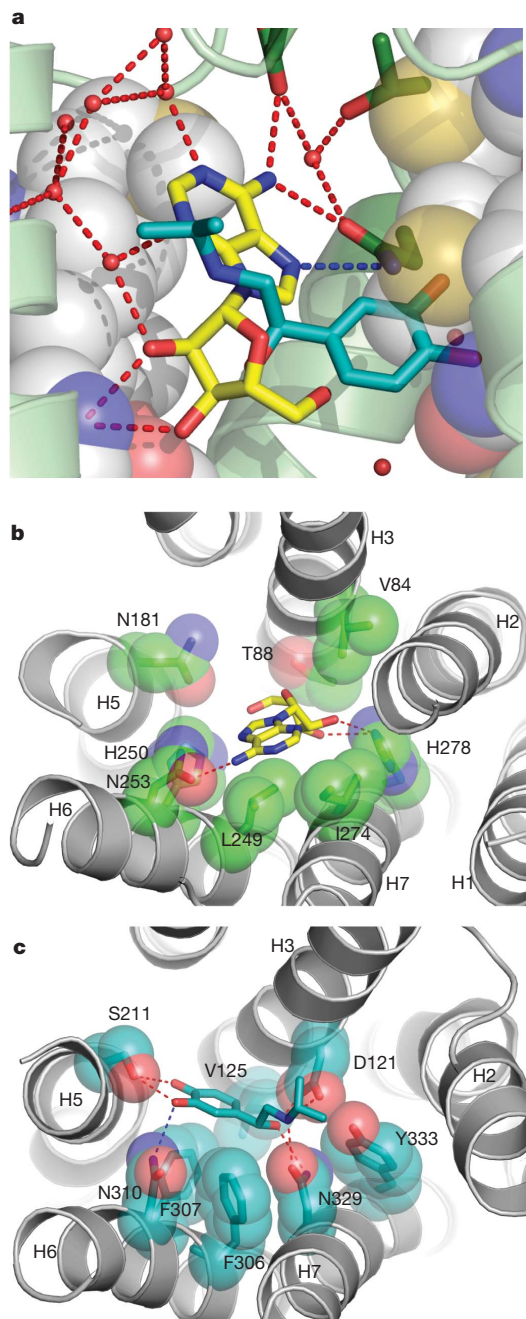


Figure 4 | Comparison of the positions of agonists in the binding pockets of the A_{2A} R and β_1 -AR. **a**, The structures of the A_{2A} R bound to adenosine and the β_1 -AR bound to isoprenaline (PDB code 2Y03)⁷ were aligned by superimposing equivalent atoms in the protein structure and the positions of both ligands are shown as stick models with the carbon atoms in blue/green (isoprenaline) or yellow (adenosine); N, blue; O, red. The A_{2A} R structure is shown, with H5 and H7 as space-filling models (C, grey; N, blue; O, red) and the remainder of the structure as a cartoon (pale green). Some water molecules are shown as red spheres, hydrogen bonds as red dashed lines and the polar contacts as blue dashed lines. The orientation of the figure is identical to that shown in Fig. 2. **b**, Structure of the A_{2A} R bound to adenosine viewed from the extracellular surface. **c**, Structure of β_1 -AR bound to isoprenaline (PDB code 2Y03)⁷ viewed from the extracellular surface. In panels **b** and **c**, equivalent side chains in the respective structures that make contacts to both isoprenaline and adenosine in their respective receptors are shown as space-filling models and they have the following Ballesteros–Weinstein numbers (amino acid side chains are shown in parentheses for the A_{2A} R and β_1 -AR, respectively): 3.32 (V84, D121); 3.36 (T88, V125); 5.42 (N181*, S211); 6.51 (L249, F306); 6.55 (N253, N310); 6.52 (H250*, F307); 7.39 (I274, N329); 7.43 (H278, Y333). An asterisk indicates residues that only make indirect contacts to the agonists via a water molecule.

interactions between the agonist and H3, combined with polar interactions involving residues in H7. Despite these differences, agonist binding to both receptors involves strong attractive non-covalent interactions that pull the extracellular ends of H3, H5 and H7 together, which is the necessary prerequisite to receptor activation.

While this manuscript was in review, a related manuscript appeared³⁰, describing the structure of the A_{2A} -T4L chimaera bound to the agonist UK432097, which is identical to NECA except for two large substituents on the adenine ring. The structure of UK432097-bound A_{2A} -T4L is very similar to the structures presented here in the transmembrane regions (r.m.s.d. 0.6 Å), although there are differences in the extracellular surface due to the bulky extensions of UK432097 interacting with the extracellular loops and the absence of density for residues 149–157. Xu *et al.*³⁰ conclude that the structure of UK432097-bound A_{2A} -T4L is in an “active state configuration”, whereas we conclude that the NECA- and adenosine-bound structures are best defined as representing an intermediate state between R and R*.

METHODS SUMMARY

Expression, purification and crystallization. The thermostabilized A_{2A} R-GL31 construct contains amino acid residues 1–316 of the human A_{2A} R, four thermostabilizing point mutations (L48A^{2,46}, A54L^{2,52}, T65A^{2,63} and Q89A^{3,37}) and the mutation N154A to remove a potential N-glycosylation site. A_{2A} R-GL31 was expressed in insect cells using the baculovirus expression system and purified in the detergent octylthioglucoside using Ni^{2+} -NTA affinity chromatography and size exclusion chromatography (see Methods). The purified receptor was crystallized in the presence of cholesteryl hemisuccinate by vapour diffusion, under conditions described in Methods.

Data collection, structure solution and refinement. Diffraction data were collected in multiple wedges (20° per wedge) from a single cryo-cooled crystal (100 K) for the GL31–NECA complex at beamline ID23-2 at the European Synchrotron Radiation Facility and from four crystals for the GL31–adenosine complex, at beamline I24 at the Diamond Light Source. The structures were solved by molecular replacement using the ZM241385-bound A_{2A} -T4L structure (PDB code 3EML)⁸ as a model (see Methods). Data collection and refinement statistics are presented in Supplementary Table 1 and omit densities for the ligands are shown in Supplementary Fig. 6.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 February; accepted 21 April 2011.

Published online 18 May 2011.

1. Fredholm, B. B. *et al.* International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and classification of adenosine receptors—an update. *Pharmacol. Rev.* **63**, 1–34 (2011).
2. Evans, B. A. *et al.* Ligand-directed signalling at β -adrenoceptors. *Br. J. Pharmacol.* **159**, 1022–1038 (2010).
3. Hofmann, K. P. *et al.* A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem. Sci.* **34**, 540–552 (2009).
4. Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363 (2009).
5. Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **469**, 175–180 (2011).
6. Rosenbaum, D. M. *et al.* Structure and function of an irreversible agonist– β_2 adrenoceptor complex. *Nature* **469**, 236–240 (2011).
7. Warne, T. *et al.* The structural basis for agonist and partial agonist action on a β_1 -adrenergic receptor. *Nature* **469**, 241–244 (2011).
8. Jaakola, V. P. *et al.* The 2.6 Å crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217 (2008).
9. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three dimensional models and computational probing of structure function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).
10. Kobilka, B. K. & Deupi, X. Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **28**, 397–406 (2007).
11. Yao, X. J. *et al.* The effect of ligand efficacy on the formation and stability of a GPCR–G protein complex. *Proc. Natl Acad. Sci. USA* **106**, 9501–9506 (2009).
12. Vauquelin, G. & Van Liefde, I. G protein-coupled receptors: a count of 1001 conformations. *Fundam. Clin. Pharmacol.* **19**, 45–56 (2005).
13. Palczewski, K. *et al.* Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745 (2000).
14. Li, J. *et al.* Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343**, 1409–1438 (2004).
15. Park, J. H. *et al.* Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183–187 (2008).
16. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).

17. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2 -adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265 (2007).
18. Rasmussen, S. G. *et al.* Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383–387 (2007).
19. Warne, T. *et al.* Structure of a β_1 -adrenergic G-protein-coupled receptor. *Nature* **454**, 486–491 (2008).
20. Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071 (2010).
21. Chien, E. Y. *et al.* Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091–1095 (2010).
22. Oldham, W. M. & Hamm, H. E. Heterotrimeric G protein activation by G-protein-coupled receptors. *Nature Rev. Mol. Cell Biol.* **9**, 60–71 (2008).
23. Murphree, L. J. *et al.* Human A_{2A} adenosine receptors: high-affinity agonist binding to receptor-G protein complexes containing G β_4 . *Mol. Pharmacol.* **61**, 455–462 (2002).
24. Serrano-Vega, M. J., Magnani, F., Shibata, Y. & Tate, C. G. Conformational thermostabilization of the β_1 -adrenergic receptor in a detergent-resistant form. *Proc. Natl Acad. Sci. USA* **105**, 877–882 (2008).
25. Magnani, F., Shibata, Y., Serrano-Vega, M. J. & Tate, C. G. Co-evolving stability and conformational homogeneity of the human adenosine A_{2A} receptor. *Proc. Natl Acad. Sci. USA* **105**, 10744–10749 (2008).
26. Shibata, Y. *et al.* Thermostabilization of the neurotensin receptor NTS1. *J. Mol. Biol.* **390**, 262–277 (2009).
27. Kim, S. K. *et al.* Modeling the adenosine receptors: comparison of the binding domains of A_{2A} agonists and antagonists. *J. Med. Chem.* **46**, 4847–4859 (2003).
28. Dal Ben, D. *et al.* Adenosine receptor modeling: what does the A_{2A} crystal structure tell us? *Curr. Top. Med. Chem.* **10**, 993–1018 (2010).
29. Wacker, D. *et al.* Conserved binding mode of human β_2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **132**, 11443–11445 (2010).
30. Xu, F. *et al.* Structure of an agonist-bound human A_{2A} adenosine receptor. *Science* (2011).
31. Pottorff, L. *et al.* Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D* **60**, 2288–2294 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by core funding from the Medical Research Council, and grants from Heptares Therapeutics Ltd and from the Biotechnology and Biological Sciences Research Council (BB/G003653/1). We would like to thank F. Magnani for technical help at the start of the project and F. Gorrec for developing the crystallization screen. We also thank the beamline staff at the European Synchrotron Radiation Facility, particularly at (beamline ID23-2; D. Flot and A. Popov), the Swiss Light Source (beamline X06SA) and at the Diamond Light Source (beamline I24; G. Evans, D. Axford and R. Owen). F. Marshall, M. Weir, M. Congreve and R. Henderson are thanked for their comments on the manuscript.

Author Contributions G.L. devised and performed receptor expression, purification, crystallization, cryo-cooling of the crystals, data collection, data processing and structure refinement. T.W. and P.C.E. helped with expression, crystal cryo-cooling and data collection. K.B. performed the radioligand binding assays and pharmacological analyses on receptor mutants in whole cells and C.J.L. was involved in data analysis and experimental design. A.G.W.L. was involved in data processing and structure refinement. Manuscript preparation was performed by G.L., A.G.W.L. and C.G.T. Overall project management was by C.G.T.

Author Information Co-ordinates and structure factors have been submitted to the PDB database under accession codes 2YDO and 2YDV for A_{2A} R-GL31 bound to adenosine or NECA, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.G.T. (cgt@mrc-lmb.cam.ac.uk).

METHODS

Expression, purification and crystallization. The human A_{2A} construct, GL31, contains four thermostabilizing point mutations (L48A^{2,46}, A54L^{2,52}, T65A^{2,63} and Q89A^{3,37}), the mutation N154A to remove the potential N-glycosylation site and a truncation at the C terminus after Ala 316 (ref. 32). A polyhistidine tag (His₁₀) was engineered at the C terminus, separated from the receptor by a TEV protease cleavage site. Baculovirus expression and membrane preparation were performed as described previously for the β_1 -AR³³.

Membranes were thawed at room temperature (20–25 °C), diluted with 25 mM HEPES pH 7.4, in the presence of protease inhibitors (Complete; Boehringer). Membranes were pre-incubated with NECA at 100 μ M for 45 min at room temperature. The receptor–ligand complexes were then solubilised by adding decylmaltoside (DM) and NaCl to give final concentrations of 1.5% and 0.3 M, respectively, stirred for 30 min (4 °C) and insoluble material removed by ultracentrifugation (120,000g, 45 min, 4 °C). All protein purification steps were performed at 4 °C. The solubilised receptor sample was filtered through a 0.22 μ m filter (Millipore) and applied to a 5 ml Ni-NTA superflow cartridge (Qiagen) pre-equilibrated with buffer (25 mM HEPES, pH 7.4, 0.1 M NaCl, 100 μ M NECA, 0.15% DM, 2.5 mM imidazole). The column was washed sequentially with the same buffer supplemented with either 10, 40 or 80 mM imidazole, and then eluted with 250 mM imidazole. The eluted receptor–ligand complex was mixed with His₆-tagged TEV protease to cleave the tag for 4–6 h, 4 °C, concentrated to 2 ml using an Amicon-ultra spin concentrator (Ultracel-50K, Millipore) and then imidazole was removed using a PD-10 column (GE Healthcare). Eluted fractions were further purified by binding the TEV and other contaminants to Ni-NTA (QIAGEN) pre-equilibrated in 25 mM HEPES pH 7.4, 0.1 M NaCl, 100 μ M NECA, 0.15% DM, 40 mM imidazole, incubating for 30 min and then collecting the flow-through. For detergent exchange into 0.35% octylthiogluco- side (OTG), the sample was concentrated using an Amicon-ultra concentrator (Ultracel-50K, Millipore), diluted tenfold in 25 mM HEPES pH 7.4, 0.1 M NaCl, 100 μ M NECA, 0.35% OTG, and concentrated again to 0.3 ml. The protein sample was applied to a Superdex 200 10/300 GL size-exclusion column pre-equilibrated in 25 mM HEPES pH 7.4, 0.1 M NaCl, 100 μ M NECA, 0.35% OTG and run at 0.5 ml min⁻¹. Eluted receptor fractions (2–2.5 ml) were concentrated to 50–60 μ l. Protein determination was performed using the amido black³⁴ assay.

Before crystallization, cholesteryl hemisuccinate (CHS) and OTG were added to 1 mg ml⁻¹ and 0.5% respectively and the protein concentration adjusted to 10–12.5 mg ml⁻¹. NECA and adenosine A_{2A}-GL31 crystal hits were obtained using a new PEG-based crystallization screen developed in house³⁵. Crystals were grown at 4 °C in 100 nl sitting drops using 0.05 M ADA NaOH, pH 6.4, 23.6% PEG 400, 4% v/v 2-propanol for the NECA complex. Crystals were cryoprotected by soaking in 0.05 M ADA NaOH, pH 6.4, 45% PEG 400. For the adenosine complex, crystals were initially grown in 0.05 M TrisHCl, pH 7.6, 9.6% PEG 200, 22.9% PEG 300. Crystals were cryoprotected by soaking in 0.05 M TrisHCl, pH 7.5, 15% PEG 200, 30% PEG 300. The crystals were mounted on Hampton CrystalCap HT loops and cryo-cooled in liquid nitrogen.

Data collection, structure solution and refinement. Diffraction data for the NECA complex were collected at the European Synchrotron Radiation Facility with a Mar 225 CCD detector on the microfocus beamline ID23-2 (wavelength, 0.8726 Å) using a 10 μ m focused beam and for the adenosine complex on beamline I24 at the Diamond Light Source with a Pilatus 6M detector and a 10 μ m microfocus beam (wavelength 0.9778 Å). The microfocus beam was essential for the location of the best diffracting parts of single crystals, as well as allowing several wedges to be collected from different positions. Images were processed with MOSFLM³⁶ and SCALA³⁷. The NECA complex was solved by molecular replacement with PHASER³⁸ using the A_{2A}-T4L structure (PDB code 3EML)⁸ as a model after removal of the coordinates for T4L, all solvent molecules and the inverse agonist ZM241384. This structure was then used as a starting model for the structure solution of the adenosine complex. Refinement and rebuilding were carried out with REFMAC5³⁹ and COOT⁴⁰, respectively. In the final models, 98.1% of residues were in the favoured region of the Ramachandran plot with one outlier for the NECA complex, and 97.7% with no outliers for the adenosine complex. Smile strings for NECA and adenosine were created using Sketcher and dictionary entries using Libcheck. Hydrogen bond assignments for the ligands were determined using HBPLUS⁴¹.

To facilitate a structural comparison between ZM241385-bound A_{2A}-T4L and the thermostabilized A_{2A}-GL31 with bound agonist, the structures were superimposed based on those residues in the region of the ligand-binding pocket that show the closest structural homology. This was achieved using the lsq_improve option of program O⁴² and an initial transformation based on residues at the C terminus of helix 6 and the N terminus of helix 7. The final superposition, based on residues 16–21 in H1, 51–70 in H2 and ECL1, 132–140 in H4 and ECL2, 142–146 in ECL2, 166–182 in ECL2 and H5 and 245–283 in H6, ECL3 and H7, gave an r.m.s.d. in C α positions of 0.66 Å for the 96 atoms and includes almost all residues involved in binding either ligand with the exception of those in H3. Using this transformation, the

adenine moiety of the agonist superimposes well with the equivalent atoms of the triazolo-triazene bicyclic ring of ZM241385 (r.m.s.d. 0.56 Å). Validation of the final refined models was carried out using Molprobity⁴³. Omit densities for the ligands are shown in Supplementary Fig. 6. All figures in the manuscript were generated using either Pymol (DeLano Scientific) or CCPmg⁴¹.

Binding of agonists and antagonist to A_{2A}R-GL31 expressed in CHO cells. Chinese hamster ovary (CHO) cells were maintained in culture in DMEM HAMS F12 media containing 10% FBS. Cells were transfected with plasmids expressing either wild-type adenosine A_{2A}R or A_{2A}R-GL31 using GeneJuice according to manufacturer's instructions (EMD Biosciences). Forty-eight hours after transfection, cells were harvested, centrifuged at 200g for 5 min at 4 °C and the pellet re-suspended in 20 mM HEPES, 10 mM EDTA buffer (pH 7.4). The membrane suspension was homogenized and centrifuged at 200g for 15 min at 4 °C. The supernatant was collected, the pellet re-suspended in 20 mM HEPES, 10 mM EDTA (pH 7.4) buffer and the solution homogenized and centrifuged as described previously⁴⁴. The collected supernatant was centrifuged for 30 min at 40,000g at 4 °C. Pellets were re-suspended in 20 mM HEPES, 0.1 mM EDTA to a protein concentration of 1 mg ml⁻¹ and stored at –80 °C.

Membranes from CHO cells transiently expressing wild-type or A_{2A}R-GL31 (10–15 μ g per well) were assessed using competition [³H]NECA binding in buffer containing 50 mM Tris-HCl (pH 7.4) as described previously⁴⁴. Inhibition curves were fitted to a four-parameter logistic equation to determine IC₅₀ values, which were converted to K_i values using K_d values determined by saturation binding and the [³H]NECA concentration of 10 nM.

G-protein-coupling activity of A_{2A}R-GL31 measured in whole cells. A_{2A}R-His₆ and A_{2A}R-GL31-His₆ (amino acid residues 1–316 of human A_{2A}R) were subcloned into plasmid pcDNA5/FRT/TO using KpnI and NotI restriction sites. Fip-in T-Rex HEK293 cells were maintained at 37 °C in a humidified atmosphere in Dulbecco's modified Eagle's medium without sodium pyruvate, supplemented with 4,500 mg l⁻¹ glucose, L-glutamine, 10% (v/v) FBS, 1% penicillin/streptomycin mixture and 10 μ g ml⁻¹ blasticidin. To generate stable cell lines, the cells were transfected with a ratio of 1:9 receptor cDNA in pcDNA5/FRT/TO vector and pOG44 vector using Genejuice as per manufacturer's instructions (EMD Biosciences). After 48 h, media were replaced with fresh medium supplemented with 200 μ g ml⁻¹ hygromycin B to select for stably expressing clones. Colonies were combined and tested for doxycycline-induced receptor expression. To induce receptor expression clones were treated with either 1 ng ml⁻¹ or 3 ng ml⁻¹ doxycycline for 16 h.

Cells were seeded at a density of 25,000 per well in a poly-L-lysine coated 96-well half area plate. Cells were induced with doxycycline (3 or 1 ng ml⁻¹) for 16 h. After 16 h media were removed and replaced with fresh media containing 100 μ M Ro-201724 and 2 U ml⁻¹ adenosine deaminase. Cells were incubated at 37 °C for 30 min before addition of varying concentrations of agonist (25 °C, 30 min). As a control cells were also incubated for 30 min (25 °C) with 10 μ M forskolin. Cells were then lysed and cAMP produced detected using the CisBio cAMP kit according to manufacturer's instructions before plates were read on a PolarStar fluorescence plate reader.

32. Lebon, G. Bennett, K. Jazayeri, A. & Tate, C. G. Thermostabilization of an agonist-bound conformation of the human adenosine A_{2A} receptor. *J. Mol. Biol.* doi:10.1016/j.jmb.2011.03.075 (in the press).
33. Warne, T., Chirmside, J. & Schertler, G. F. Expression and purification of truncated, non-glycosylated turkey β -adrenergic receptors for crystallization. *Biochim. Biophys. Acta* **1610**, 133–140 (2003).
34. Schaffner, W. & Weissmann, C. A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal. Biochem.* **56**, 502–514 (1973).
35. Gorrec, F., Palmer, C., Lebon, G. & Warne, T. Pi sampling: a methodical and flexible approach to macromolecular crystallization initial screening. *Acta Crystallogr. D* **67**, 463–470 (2011).
36. Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **62**, 48–57 (2006).
37. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82 (2006).
38. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
39. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
41. McDonald, I. K. & Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793 (1994).
42. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
43. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
44. Robertson, N. et al. The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. *Neuropharmacology* **60**, 36–44 (2011).

Principles of activation and permeation in an anion-selective Cys-loop receptor

Ryan E. Hibbs¹ & Eric Gouaux^{1,2}

Fast inhibitory neurotransmission is essential for nervous system function and is mediated by binding of inhibitory neurotransmitters to receptors of the Cys-loop family embedded in the membranes of neurons. Neurotransmitter binding triggers a conformational change in the receptor, opening an intrinsic chloride channel and thereby dampening neuronal excitability. Here we present the first three-dimensional structure, to our knowledge, of an inhibitory anion-selective Cys-loop receptor, the homopentameric *Caenorhabditis elegans* glutamate-gated chloride channel α (GluCl), at 3.3 Å resolution. The X-ray structure of the GluCl–Fab complex was determined with the allosteric agonist ivermectin and in additional structures with the endogenous neurotransmitter L-glutamate and the open-channel blocker picrotoxin. Ivermectin, used to treat river blindness, binds in the transmembrane domain of the receptor and stabilizes an open-pore conformation. Glutamate binds in the classical agonist site at subunit interfaces, and picrotoxin directly occludes the pore near its cytosolic base. GluCl provides a framework for understanding mechanisms of fast inhibitory neurotransmission and allosteric modulation of Cys-loop receptors.

Fast inhibitory neurotransmission modulates both the magnitude and duration of neuronal activity, occurs on a timescale of milliseconds, and involves the release of inhibitory neurotransmitters into the synapse and activation of the cognate ligand-gated ion channels. As demonstrated nearly 60 years ago¹, fast inhibitory neurotransmission leads to an increase in the permeability of the cell membrane to chloride, the most abundant biological anion. Because the membrane potential at which chloride is at equilibrium is near the neuronal resting potential, neurotransmitter-gated, chloride-selective ion channels generally oppose normal excitability and repolarize the cell².

The neurotransmitter receptors that directly mediate chloride permeability constitute one half of the Cys-loop receptor family³. Receptors in this family are composed of five either identical or homologous subunits, which generate diversity in functional profiles and pharmacological preferences. Cys-loop receptors fall into two broad categories. The cation-selective members are the nicotinic acetylcholine (nAChR) and serotonin 5-HT₃ receptors. Those selective for anions include the γ -aminobutyric acid (GABA_{A/C}), glycine receptors and invertebrate glutamate-gated chloride channels (GluCl)^{3–6}. So far, there is no structural information for an anion-selective Cys-loop receptor, and the mechanism by which chloride is selected remains unclear.

Ligand-gated chloride channels are critical not only for maintaining appropriate neuronal activity, but have long been important therapeutic targets: benzodiazepines, barbiturates, some intravenous and volatile anaesthetics, alcohol, strychnine, picrotoxin and ivermectin all derive their biological activity from acting on the inhibitory half of the Cys-loop receptor family^{3,7}. Of note is that many of the therapeutically useful compounds acting at Cys-loop receptors target an allosteric site. The sites in Cys-loop receptors at which these allosteric ligands bind and their structure-based mechanisms of action are largely unresolved.

Crystallization of GluCl–Fab complex

We identified the *Caenorhabditis elegans* GluCl α glutamate-gated chloride channel⁸ as a promising candidate using fluorescence-detection size-exclusion chromatography (FSEC)⁹. In comparison to human

Cys-loop receptors, GluCl α is most similar to the α 1 glycine receptor, with which it shares 34% amino acid sequence identity (see alignment in Supplementary Fig. 1). Optimization of the receptor construct for crystallization (GluCl_{cryst}) was guided by FSEC analysis and required deletion of 41 residues from the amino terminus, 6 residues from the carboxy terminus and replacement of the M3–M4 loop (Lys 345–Lys 402) with an Ala–Gly–Thr tripeptide. Well-ordered crystals diffracting to \sim 3.3 Å resolution required co-crystallization of GluCl_{cryst} as a complex with a Fab, ivermectin and lipids (Supplementary Fig. 2). Structures with agonist or channel blocker at 3.3 and 3.4 Å were obtained by soaking GluCl_{cryst}–Fab–ivermectin crystals with glutamate or picrotoxin, respectively. The electron density maps are of high quality, thus enabling the positioning of almost all receptor residues and refinement to satisfactory crystallographic residuals and stereochemistry (Supplementary Table 1 and Supplementary Fig. 3).

Architecture

The GluCl_{cryst}–Fab complex forms a pinwheel shape comprising a cylindrical homopentamer of GluCl_{cryst} subunits with Fab molecules bound at each subunit interface (Fig. 1a, b). Each GluCl_{cryst} subunit consists of a large N-terminal extracellular domain of mostly β structure, followed by four α -helical transmembrane spans (M1–M4; Fig. 1c). The overall architecture of the extracellular domain is similar to that found in the bacterial receptor orthologues from *Gloeobacter violaceus* (GLIC)^{10,11} and *Erwinia chrysanthemi* (ELIC)¹². There is an additional helix at the N terminus reminiscent of the acetylcholine-binding protein^{13–15} (AChBP) and *Torpedo marmorata* nAChR¹⁶ structures. Significantly, GluCl contains the Cys-loop disulphide strictly conserved in eukaryotes as well as a disulphide bond in loop C present in glycine receptors (Fig. 1c). The transmembrane helices adopt a fold like the bacterial receptors and nAChR, with the five M2 segments lining the pore and adopting an open channel conformation, akin to the conformation visualized in the GLIC structures.

To understand the molecular principles of ion channel activation, agonist binding and ion channel permeation and block, we determined

¹Vollum Institute, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA. ²Howard Hughes Medical Institute, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA.

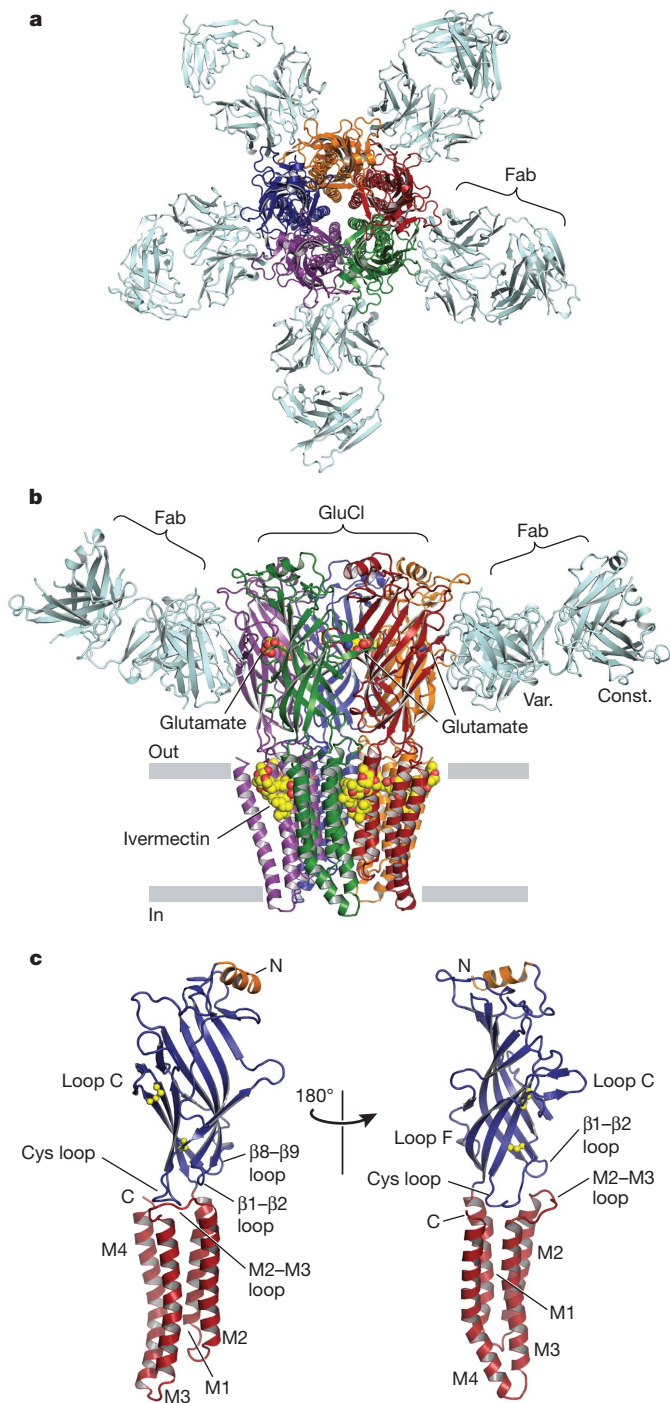


Figure 1 | Architecture of the GluCl_{cryst}-Fab complex. **a**, View of the GluCl_{cryst}-Fab complex looking down pore axis towards cytosol. Fab molecules (cyan) are bound at each GluCl_{cryst} subunit interface. **b**, View parallel to lipid membrane; only two Fab molecules are shown for clarity. The ligands ivermectin and glutamate are represented as spheres with carbon atoms in yellow, oxygen in red and nitrogen in blue. Const., constant region; Var., variable region. **c**, A single GluCl_{cryst} subunit from two angles, approximate orientation as in panel **b**. The Cys loop and loop C disulphide bonds are shown as spheres, N and C termini and transmembrane spans are indicated. Loops of particular relevance to agonist binding and allosteric gating linkage are also indicated.

separate crystal structures with the allosteric agonist ivermectin, and with ivermectin and glutamate, picrotoxin or iodide. Ivermectin is bound at each of the GluCl_{cryst} subunit interfaces in the transmembrane domain whereas glutamate electron density is present in all five of the classical neurotransmitter-binding sites in the extracellular

domain. Anomalous difference density for iodide is present in sites at the base of the transmembrane pore in a region important for ion selectivity, and a chloride ion was fit into non-protein electron density in the ion channel pore adjacent to the binding site for picrotoxin.

Allosteric activation and modulation

Ivermectin is a semi-synthetic macrocyclic lactone and broad-spectrum antiparasitic agent, widely used to treat river blindness in humans and parasitic infections in animals^{17,18}. It achieves its margin of therapeutic efficacy by activating invertebrate glutamate-gated chloride channels at nanomolar concentrations^{8,19}, yet it also manifests activating and potentiating activities on vertebrate Cys-loop receptors^{20–22} and on P2X ATP-gated ion channels²³ at higher concentrations. Ivermectin potentially activates GluCl α (Supplementary Fig. 4) while simultaneously rendering the receptor susceptible to further activation by glutamate²⁴. Hence, at GluCl α , we deem ivermectin a partial allosteric agonist.

Ivermectin binds at subunit interfaces on the periphery of the transmembrane domains, proximal to the extracellular side of the membrane bilayer (Fig. 2a, b and Supplementary Figs 5–7). Wedged between the M3 α -helix on the principal or (+) subunit and the M1 α -helix on the complementary or (–) subunit, ivermectin inserts deeply into the subunit interface and makes important contacts with the M2 (+) pore-lining α helix and the M2–M3 loop. Its site occupies approximately two turns of helix on the M1 and M3 helices and centres on a single turn of π helix between residues Leu 217 and Ile 222 on M1, as illustrated by a hydrogen bond between the main-chain carbonyl oxygen of Leu 218 and a tertiary hydroxyl on ivermectin (Fig. 2c). Through extensive hydrophobic interactions and one hydrogen bond with each of the M1, M2 and M3 α -helices, ivermectin buries 278 and 254 Å² of surface area on the (+) and (–) subunits in the interface.

Ser 260 forms a hydrogen bond with the secondary hydroxyl group on the deeply buried cyclohexene ring of ivermectin (Fig. 2a–c and Supplementary Fig. 8). A serine residue in this position is correlated with direct activation by ivermectin in other Cys-loop receptors. Glycine and GABA_A receptors have a serine in the equivalent position and are directly activated by ivermectin^{20,21}, yet there is no similar serine in $\alpha 7$ nAChRs, where ivermectin is a positive allosteric modulator but does not directly activate²², nor in GluCl β receptors, where ivermectin has no activity⁸. The equivalent position is critical for GABA_A receptor modulation by alcohol²⁵, anticonvulsants, anaesthetics and diuretics; glycine and 5-HT₃ receptor modulation by anaesthetics⁶; and $\alpha 7$ nAChR modulation by additional compounds²⁶. Hence, the ivermectin binding site in GluCl_{cryst} is shared, at least in part, by many important modulators of Cys-loop receptors. In GluCl we suspect that the interaction of ivermectin with the pore-lining M2 helix increases both its affinity for the receptor and its ability to stabilize the open state.

Ivermectin binding to GluCl probably results in two types of conformational changes: first, a local distortion of the receptor in the vicinity of the binding site; and second, a global conformational change of the receptor that corresponds to a transition from a closed, resting state to an open, activated state. Because we lack a structure of GluCl_{cryst} in the absence of ivermectin, GLIC provides a reference for gauging the local structural consequences of ivermectin binding to the transmembrane domain of the receptor. In comparing these two structures we find that the binding of ivermectin increases the separation between M1 and M3 of adjacent subunits, as defined by a 9.4 Å spacing between GluCl_{cryst} Leu 218 and Gly 281 C α atoms compared to a 6.4 Å spacing for the corresponding atoms in GLIC. This splaying apart of the transmembrane helices in GluCl_{cryst} occurs at the level of a strictly conserved proline residue in M1 that forms the C-terminal end of the short π helix (Supplementary Fig. 9).

We hypothesize that the global conformational change induced by ivermectin binding is rooted in the splaying apart of the M1 and M3 helices and the movement of the apical portion of M2 away from the

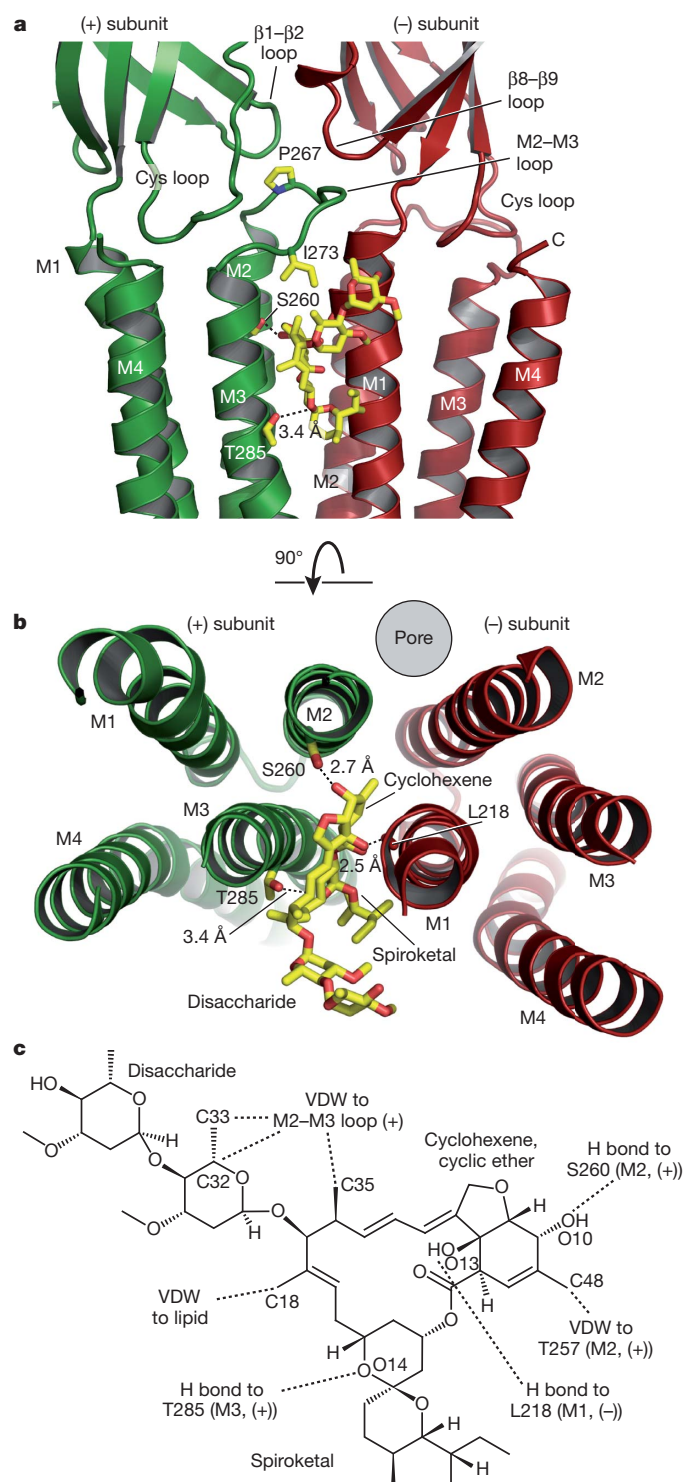


Figure 2 | Ivermectin-binding site and atomic interactions. **a, b**, Two orientations of a GluCl subunit interface focusing on ivermectin-binding site. Dashed lines indicate hydrogen bonds. In **a**, view is from receptor periphery looking parallel to the membrane, and in **b** looking down pore from extracellular side with the extracellular domain removed for clarity. **c**, Chemical structure of ivermectin with interactions indicated. VDW, van der Waals. Atomic numbering is from PDB file.

pore axis, towards the periphery of the receptor, opening an ion conductive pathway. This open-pore conformation of M2 is then stabilized through interactions between ivermectin and the apical end of M2. In addition, ivermectin may stabilize the open state of the ion channel through contacts between the disaccharide moiety and Ile 273 in the M2–M3 loop (Supplementary Fig. 10).

Neurotransmitter-binding site

The ion channel of GluCl_{cryst} is activated by glutamate only after activation by ivermectin (Supplementary Fig. 11), in a manner similar to full-length GluCl α (ref. 24). The homomeric GluCl β and heteromeric GluCl $\alpha\beta$ receptors, by contrast, are directly activated by glutamate⁸. In the context of GluCl_{cryst}, micromolar concentrations of glutamate augment ivermectin-induced currents by 30–70%, similar to that of full-length receptor. [³H]-L-glutamate binds directly to the GluCl_{cryst} receptor with a dissociation constant (K_d) of 680 nM (Supplementary Figs 11–13). In agreement with the electrophysiology experiments, [³H]-L-glutamate binding requires ivermectin. To understand the molecular basis of glutamate binding we determined the structure of GluCl_{cryst} in the presence of ivermectin and glutamate. Sausage-shaped electron density assigned to glutamate was $\sim 8\sigma$ in $F_o - F_c$ omit maps in all five of the classical agonist-binding sites. Omit electron density maps subjected to real space, five-fold averaging showed a protrusion in the electron density ‘sausage’ that we attributed to the α -amino group of glutamate (Supplementary Fig. 14).

Glutamate binds in the classical neurotransmitter site in the extracellular domain^{3,27}, lodged between subunits and nearly inaccessible to solvent (Fig. 3a, b). The architecture of the site is box-like, with loops from the (+) subunit forming ‘sides’ of the binding site and the β strands on the (–) subunit defining the ‘base.’ Loop C, postulated to have a critical role in allosteric activation^{15,28–30}, adopts a closed conformation consistent with AChBP structures bound by agonists. Functional groups on glutamate bridge the (+) and (–) subunits with the α -substituents snugly sandwiched between Tyr 151 and Tyr 200 on the (+) subunit, and positively charged residues, including Arg 37 from a region important for conotoxin–nAChR interaction³¹, and Arg 56 on the (–) subunit, making contacts with the α - and γ -carboxylate groups. These arginine residues, in combination with neighbouring cationic amino acids, provide the binding pocket with a strongly positive electrostatic potential (Supplementary Fig. 15). The α -amino nitrogen of glutamate is stabilized through a 3.8 Å cation– π interaction with Tyr 200 on loop C, a hydrogen bond with the backbone carbonyl oxygen of Ser 150 and a close interaction with the backbone carbonyl oxygen of Tyr 151. A comparison of the determinants of glutamate binding with the corresponding residues in the AChBPs and other receptors is made in Supplementary Fig. 16.

To test the sensitivity of the glutamate binding site to perturbations in ligand structure, we screened glutamate analogues for competition with [³H]-L-glutamate bound to the ivermectin-complexed receptor (Fig. 3c and Supplementary Fig. 17). L-Glutamate bound much tighter than L-homocysteine sulphonic acid, which differs only in replacement of C δ with sulphur. Extending the side-chain length with an extra carbon (L-amino adipic acid) or shortening it (L-aspartate) resulted in a further drop in affinity, and changing the stereochemistry (D-glutamate) or removing the side-chain negative charge but not its ability to hydrogen bond (L-glutamine) decreased binding further. Thus, the GluCl neurotransmitter-binding pocket is selective for small dicarboxylate L-amino acids, consistent with the constellation of atomic interactions between agonist and receptor (Fig. 3a, b).

Upon the binding of glutamate the side chain of Arg 56 in loop D (β 2) shifts by ~ 0.5 Å to accommodate the agonist and Tyr 200 in loop C repositions by ~ 0.5 Å closer to the ligand, small yet significant conformational changes consistent with movements of loops C and D in the agonist-induced activation of the receptor. These residues, together with Arg 37 (β 1), are located on elements of protein structure directly connected to the ion channel pore. We suggest that ivermectin, a partial allosteric agonist, stabilizes an ‘activated’ conformation of the agonist site and that binding of glutamate to this ‘activated’ site further stabilizes the open state of the receptor, increasing chloride conductance. Ivermectin may transduce a conformational change to the neurotransmitter site through its interactions with the M2–M3 loop, located at the structural nexus of three extracellular domain loops central to allosteric communication between the neurotransmitter site in the extracellular domain and the transmembrane pore: the Cys,

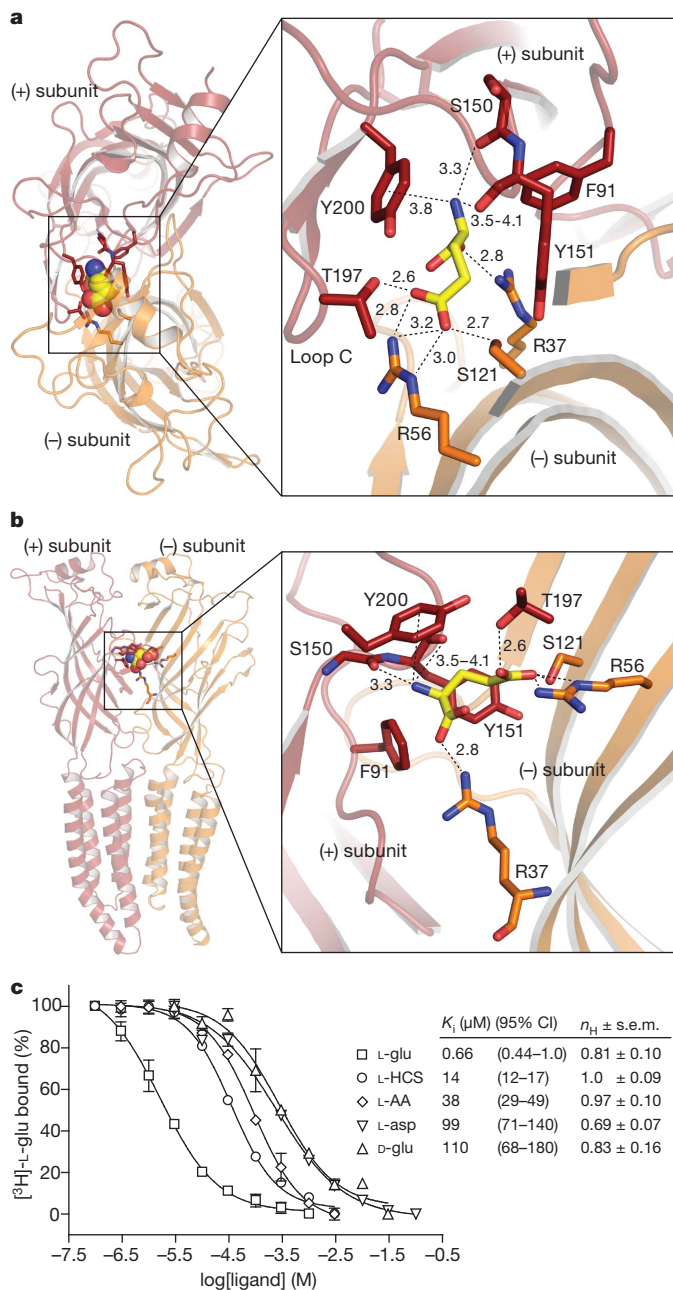


Figure 3 | Glutamate-binding site and specificity. **a**, View from extracellular side towards membrane at glutamate in binding site in subunit interface. **b**, View of binding site looking parallel to membrane with loop C removed for clarity. Dashed lines with distances in Å indicate hydrogen bonding and, in the case of Tyr 200, cation- π interactions. Unless a range is given, distances are an average from the five binding sites. **c**, Radioligand competition experiments with L-glutamate and congeners against 1 mM [3 H]-L-glutamate. Calculated inhibition constant (K_i) values assume a K_d for [3 H]-L-glutamate of 680 nM and are shown in inset table. $n = 2$. CI, confidence interval. L-HCS and L-AA are L-homocysteine sulphonic acid and L-amino adipic acid, respectively. Error bars are s.e.m. and n_H is the Hill coefficient.

$\beta 1$ – $\beta 2$ and $\beta 8$ – $\beta 9$ loops⁶ (Fig. 2a and Supplementary Fig. 10). Hydrophobic residues in the M2–M3 loop mediating these interactions are well conserved in most Cys-loop receptors, consistent with the M2–M3 loop having a central role in the activation mechanism of receptors throughout the family^{32–36}.

Pore conformation

To test the hypothesis that the GluCl_{cryst}–ivermectin structure represents an open, conducting conformation (Fig. 4), we carried out

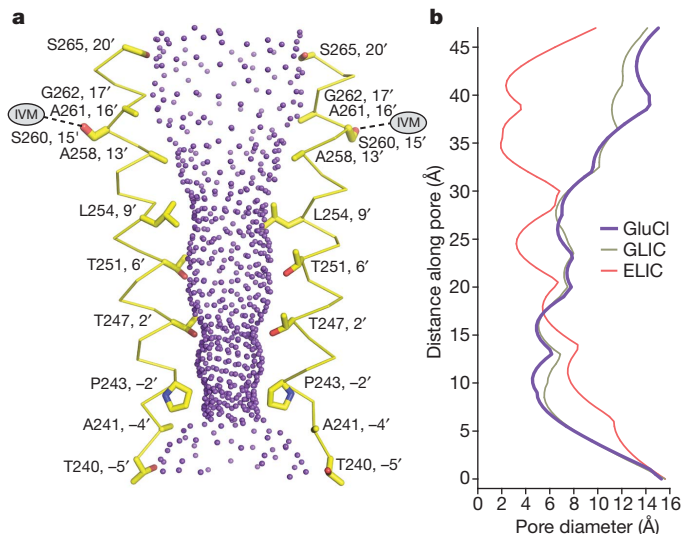


Figure 4 | Ion channel. **a**, Purple spheres represent internal surface of transmembrane ion channel, with side chains shown for pore-lining residues from two of the five M2 α helices that line the pore; Ser 260 does not line the pore but hydrogen bonds with ivermectin (IVM). **b**, Pore diameter is plotted as a function of longitudinal distance along the pore for GluCl_{cryst} open (GLIC; PDB code 3EAM) and closed (ELIC; PDB code 2VL0) bacterial receptors.

functional and structural studies using picrotoxin, an open channel blocker^{37,38} (Fig. 5 and Supplementary Fig. 18). Electron density in picrotoxin-soaked crystals was apparent at a position near the cytosolic side of the transmembrane pore, on the five-fold axis of molecular symmetry (4.3σ ; Supplementary Fig. 19). Thus, the observed electron density is an average of five orientations. Nevertheless, the egg-shaped picrotoxin-associated electron density indicates that the basket-like, fused tricyclic rings are directed extracellularly and near the 2' Thr, whereas the isoprenyl tail points towards the cytoplasm and is proximal to the -2' Pro residues. In this position, the majority of the oxygen atoms of picrotoxin are cradled by the polar belt of 2' Thr hydroxyls whereas the hydrophobic isoprenyl moiety is surrounded by the methylene groups of the non-polar -2' Pro side chains. Most importantly, the binding of picrotoxin to the pore of the GluCl_{cryst}–ivermectin complex reinforces our hypothesis that the pore is in an open conformation.

The smallest diameter of the GluCl_{cryst} ion channel pore is ~ 4.6 Å, defined by a hydrophobic 'girdle' of -2' Pro side chains proximal to the cytoplasmic side of the membrane. Because chloride has a Pauling

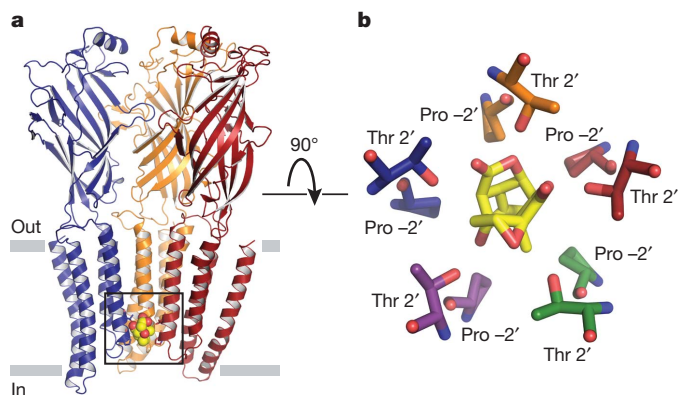


Figure 5 | Picrotoxin-binding site. **a**, The front two subunits have been removed to show the picrotoxin location (boxed) at the cytosolic base of the pore. Residues involved in picrotoxin binding are shown as sticks and van der Waals surfaces are shown for picrotoxin. **b**, View looking into pore from the extracellular domain at the picrotoxin position relative to the 2' Thr and -2' Pro side chains. Picrotoxin is shown as sticks with carbon atoms in yellow and oxygen atoms in red.

radius of 1.8 Å (ref. 2), passage of chloride, iodide (Pauling radius of 2.2 Å (ref. 2)) and other permeant ions through the $-2'$ Pro constriction must involve substantial dehydration, in agreement with previous studies demonstrating a correlation between energies of hydration and relative permeabilities (higher for iodide than chloride; Supplementary Fig. 20)³⁹. The pore constriction in GluCl_{cryst} is somewhat smaller than that estimated for GABA_A, glycine and GluCl β receptors (5.2–6.2 Å)^{3,40}, based on low but measureable relative permeability to ions like acetate, gluconate and phosphate. This difference may be due to the alanine residues at the $-2'$ position in the β subunits of all three of those receptors.

GluCl is related to the *Torpedo* nAChR (PDB code 2BG9)¹⁶ in amino acid sequence and three-dimensional structure and thus we compared the structures and the aligned sequences. In so doing, we found inconsistencies between amino acid sequence-based alignments and three-dimensional structure-based alignments of the M2 and M3 α helices. A similar finding was described in comparisons of GLIC to the nAChR⁴. Our analysis indicates that in the α subunit M2 pore-lining helix and the M3 α -helix, the nAChR amino acid assignment is off in register by 4 residues or ~ 1 turn of an α -helix beginning with the M1–M2 loop (Supplementary Fig. 21).

Ion selectivity

Analysis of GluCl_{cryst} surface electrostatics reveals an electropositive vestibule, a slightly electronegative extracellular half of the transmembrane pore, and an electropositive intracellular half (Fig. 6a). None of the pore-lining residues in GluCl_{cryst} bear a formal charge and thus the positive electrostatic potential at the base of the pore arises from the oriented peptide dipoles in the M2 α helices⁴¹, reminiscent of the role that helical dipoles have in ClC chloride channels⁴². Cation channels

reverse the selectivity imposed by orientation of the M2 dipoles through placing a negatively charged side chain near the pore constriction point⁴³. Although other regions contribute to the modulation of conductance and selectivity in some Cys-loop receptors, the minimal determinants of selectivity are the $-1'$ Ala and $-2'$ Pro positions for anions and the $-1'$ Glu for cations^{3,44}, with no requirement for positively charged amino acids in the pore of anion-selective channels⁴⁰ (Supplementary Fig. 22).

To identify sites important in chloride binding and selectivity, we soaked crystals of GluCl_{cryst} in iodide, a heavy atom analogue of chloride, and observed four anomalous difference peaks that we ascribe to iodide, located at the cytosolic base of the transmembrane pore and centred around the five-fold symmetry axis (Fig. 6b–d). The weak density at the fifth site is simply the consequence of an interfering lattice contact with an adjacent Fab. Each iodide sits in a concave pocket of positive electrostatic potential formed by $-2'$ Pro residues from the M2 helices of adjacent subunits, main-chain backbone atoms of $-1'$ Ala and $-3'$ Ile and the methyl group of $-1'$ Ala. All three of these residues are important in selectivity for some receptors, with the $-1'$ position being an essential component of selectivity across the family³. Previous studies indicate that the main chain amide nitrogen at the $-3'$ position is important in GluCl β receptors for anion dehydration⁴⁰. In GluCl_{cryst} this atom is ~ 5 Å from the centre of the iodide anomalous density and could form water-mediated hydrogen bonds to anions at the mouth of the ion channel pore.

Electron density maps derived from all GluCl_{cryst} X-ray diffraction data sets exhibit a spherical peak in the pore between the $2'$ Thr and $6'$ Thr residues (6.8σ in $F_o - F_c$ omit maps) with no other peaks in the pore above 2.5σ . Anomalous difference electron density maps were inconclusive in identification of this peak. Therefore, we placed

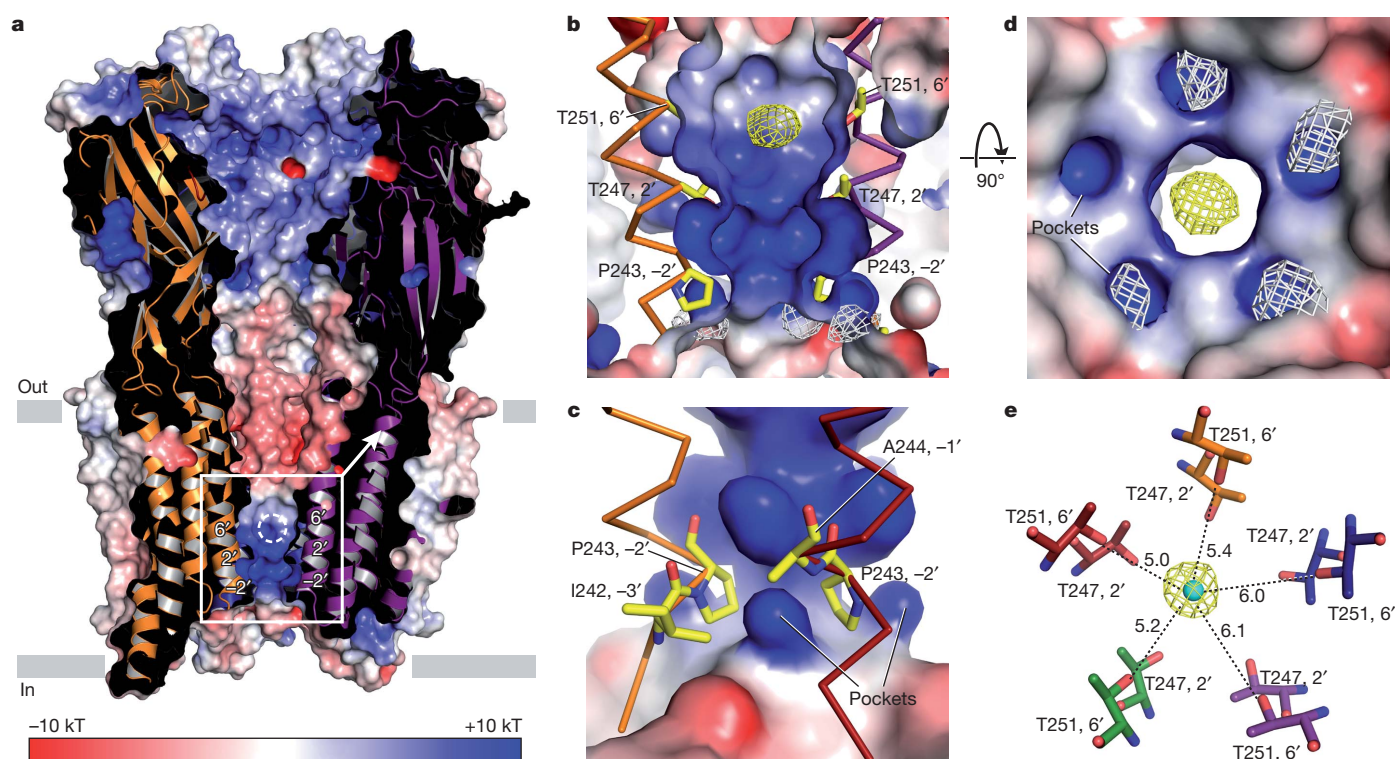


Figure 6 | Ion selectivity. **a**, The front of the receptor is cut away to reveal the interior surface of the pore, coloured by electrostatic potential. Dashed circle in the pore indicates putative chloride-binding site. Boxed area is expanded in **b**. **b**, Expanded view from **a** showing selected M2 side chains from opposing subunits. Anomalous difference peaks at pore base are attributed to iodide-binding sites (light grey mesh, contoured at 3.5σ). $F_o - F_c$ omit density for the putative chloride site is represented by yellow mesh contoured at 3σ . **c**, Shown are the electropositive pockets where iodide ions bind, viewed from inside the

protein surface; four residues from adjacent M2 helices that coordinate the iodide sites are shown as sticks. **d**, Electropositive pockets viewed from the intracellular side. **e**, Putative chloride site viewed from the extracellular side of the pore with $6'$ Thr residues in the foreground; $F_o - F_c$ omit density (yellow) is contoured at 4σ . Carbon atoms are coloured by chain and chloride is represented by a 1 Å cyan sphere. Closest distances in Å from protein atoms to centre of sphere are indicated by dashed lines from the $6'$ Thr side-chain hydroxyl.

several different ions or molecules in the difference density and determined, through crystallographic refinement and inspection of difference electron density maps, that a single chloride anion best accounted for this electron density feature (Supplementary Table 1). Distances between the modelled chloride and 6' Thr side-chain hydroxyl oxygen atoms are consistent with water-mediated hydrogen bonding of chloride in the pore⁴⁵, indicating that this location could be a transiently occupied ion-binding site flanking the constriction point. Further experimentation is required to validate the chemical identity of the bound species.

The iodide-binding sites nestled in electropositive pockets at the base of the pore suggest general principles of ion selectivity in Cys-loop receptors. In GluCl and other chloride-selective receptors there is either an alanine or glycine residue at the −1' position of the M2 helix, thus preserving the concave pocket. By contrast, in eukaryotic cation-selective channels, the −1' residue is a conserved glutamate. We suggest that the carboxylate side chain of glutamate not only fills the 'anion pocket' but that it also imposes a local negative electrostatic potential important for cation selectivity (Supplementary Fig. 22). Previous cysteine accessibility studies in cation-selective channels have indicated that the −1' Glu position lines the transmembrane pore⁴⁶. However, on the basis of the GluCl_{cryst} structure and amino acid sequence alignments, we propose that the preceding residue, a conserved glycine (−2' residue), lines the pore of cation channels, consistent with the significantly larger pore diameter of cation channels (7.4–8.4 Å). In support of the −2' residue defining the pore constriction, deletion of the −2' Pro in glycine receptors, which would shift the following glycine residue into the −2' position, increases pore diameter to 6.9 Å (ref. 47). Furthermore, the −2' Gly position in cation-selective 5-HT_{3A} receptors is accessible to modification when the pore is open⁴⁸. We propose that the −2' position lines the pore in both anion and cation channels and that the 'anion pockets' in GluCl_{cryst} are important determinants of ion selectivity, increasing the local concentration of anions at the cytoplasmic mouth of the pore.

Conclusion

Here we present the first X-ray structure, to our knowledge, of a eukaryotic Cys-loop receptor, a glutamate-gated chloride channel from *C. elegans*. GluCl_{cryst} was co-crystallized with ivermectin, a partial allosteric agonist that sequesters within the membrane bilayer and binds to exposed sites on the transmembrane domains of the receptor. Lipophilic modulators of other Cys-loop receptors may exploit a similar mechanism of interaction, including the neurosteroids at the GABA_A receptor⁶ and cholesterol at the muscle nAChR⁴⁹. The GluCl_{cryst}–ivermectin structure maps a previously uncharacterized binding site at a protein–lipid interface and defines a protein/chemical scaffold for design of receptors and ligands with new pharmacological properties and receptor specificities. Binding of ivermectin induces local changes in the membrane domain and global conformational changes in the entire receptor, pre-organizing the agonist binding site ~30 Å away and opening the ion channel pore. Analysis of amino acids lining and proximal to the pore indicates that anion selectivity is accomplished largely through a pore constriction imposed by proline residues and a positive electrostatic potential, conferred by the N-terminal end of the M2 helix dipoles. These new findings advance our understanding of the molecular mechanism of fast neuronal inhibition, the importance of which was first appreciated more than one hundred years ago⁵⁰.

METHODS SUMMARY

GluCl_{cryst} was expressed from baculovirus-infected Sf9 cells and purified by metal ion affinity chromatography. The Fab complex was isolated by size-exclusion chromatography. The GluCl_{cryst}–Fab complex was concentrated to 1–2 mg ml^{−1} and supplemented with synthetic lipids and ivermectin. Crystallization was performed by hanging-drop vapour diffusion at 4 °C with a precipitating solution containing 21–23% PEG 400, 50 mM sodium citrate pH 4.5 and 70 mM sodium

chloride. Cryoprotection was achieved by soaking crystals in precipitant solution supplemented with 30% PEG 400. Additional complexes were obtained by soaking crystals in cryoprotectant containing L-glutamate, picrotoxin or sodium iodide. Diffraction data were indexed, integrated and scaled and the structure solved by molecular replacement using a GLIC-derived homology model of GluCl_{cryst} and a Fab homology model as search probes. The molecular replacement phases were used to initiate autobuilding and the resulting model was iteratively improved by cycles of manual adjustment and crystallographic refinement. Function of GluCl was examined by two-electrode voltage clamp experiments and by [³H]-L-glutamate saturation and competition binding assays.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 March; accepted 21 April 2011.

Published online 15 May 2011.

- Coombs, J. S., Eccles, J. C. & Fatt, P. The specific ionic conductances and the ionic movements across the motoneuronal membrane that produce the inhibitory post-synaptic potential. *J. Physiol. (Lond.)* **130**, 326–374 (1955).
- Hille, B. *Ion Channels of Excitable Membranes* (Sinauer Associates, 2001).
- Thompson, A. J., Lester, H. A. & Lummis, S. C. The structural basis of function in Cys-loop receptors. *Q. Rev. Biophys.* **43**, 449–499 (2010).
- Corring, P. J. et al. Atomic structure and dynamics of pentameric ligand-gated ion channels: new insight from bacterial homologues. *J. Physiol. (Lond.)* **588**, 565–572 (2010).
- Hilf, R. J. & Dutzler, R. A prokaryotic perspective on pentameric ligand-gated ion channel structure. *Curr. Opin. Struct. Biol.* **19**, 418–424 (2009).
- Miller, P. S. & Smart, T. G. Binding, activation and modulation of Cys-loop receptors. *Trends Pharmacol. Sci.* **31**, 161–174 (2010).
- Garcia, P. S., Kolesky, S. E. & Jenkins, A. General anesthetic actions on GABA_A receptors. *Curr. Neuropharmacol.* **8**, 2–9 (2010).
- Cully, D. F. et al. Cloning of an avermectin-sensitive glutamate-gated chloride channel from *Caenorhabditis elegans*. *Nature* **371**, 707–711 (1994).
- Kawate, T. & Gouaux, E. Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **14**, 673–681 (2006).
- Hilf, R. J. & Dutzler, R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457**, 115–118 (2009).
- Bocquet, N. et al. X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* **457**, 111–114 (2009).
- Hilf, R. J. & Dutzler, R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452**, 375–379 (2008).
- Brejck, K. et al. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411**, 269–276 (2001).
- Celie, P. H. et al. Crystal structure of acetylcholine-binding protein from *Bulinus truncatus* reveals the conserved structural scaffold and sites of variation in nicotinic acetylcholine receptors. *J. Biol. Chem.* **280**, 26457–26466 (2005).
- Hansen, S. B. et al. Structures of *Aplysia* AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *EMBO J.* **24**, 3635–3646 (2005).
- Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4 Å resolution. *J. Mol. Biol.* **346**, 967–989 (2005).
- Campbell, W. C., Fisher, M. H., Stapley, E. O., Albers-Schonberg, G. & Jacob, T. A. Ivermectin: a potent new antiparasitic agent. *Science* **221**, 823–828 (1983).
- Aziz, M. A., Diallo, S., Diop, I. M., Lariviere, M. & Porta, M. Efficacy and tolerance of ivermectin in human onchocerciasis. *Lancet* **320**, 171–173 (1982).
- Arena, J. P., Liu, K. K., Paress, P. S. & Cully, D. F. Avermectin-sensitive chloride currents induced by *Caenorhabditis elegans* RNA in *Xenopus* oocytes. *Mol. Pharmacol.* **40**, 368–374 (1991).
- Adelsberger, H., Lepier, A. & Dudel, J. Activation of rat recombinant $\alpha 1\beta 2\gamma 2\delta$ GABA_A receptor by the insecticide ivermectin. *Eur. J. Pharmacol.* **394**, 163–170 (2000).
- Shan, Q., Haddrill, J. L. & Lynch, J. W. Ivermectin, an unconventional agonist of the glycine receptor chloride channel. *J. Biol. Chem.* **276**, 12556–12564 (2001).
- Krause, R. M. et al. Ivermectin: a positive allosteric effector of the $\alpha 7$ neuronal nicotinic acetylcholine receptor. *Mol. Pharmacol.* **53**, 283–294 (1998).
- Silberberg, S. D., Li, M. & Swartz, K. J. Ivermectin interaction with transmembrane helices reveals widespread rearrangements during opening of P2X receptor channels. *Neuron* **54**, 263–274 (2007).
- Etter, A., Cully, D. F., Schaeffer, J. M., Liu, K. K. & Arena, J. P. An amino acid substitution in the pore region of a glutamate-gated chloride channel enables the coupling of ligand binding to channel gating. *J. Biol. Chem.* **271**, 16035–16039 (1996).
- Ueno, S., Wick, M. J., Ye, Q., Harrison, N. L. & Harris, R. A. Subunit mutations affect ethanol actions on GABA_A receptors expressed in *Xenopus* oocytes. *Br. J. Pharmacol.* **127**, 377–382 (1999).
- Young, G. T., Zwart, R., Walker, A. S., Sher, E. & Millar, N. S. Potentiation of $\alpha 7$ nicotinic acetylcholine receptors via an allosteric transmembrane site. *Proc. Natl Acad. Sci. USA* **105**, 14686–14691 (2008).
- Kao, P. N. et al. Identification of the α subunit half-cystine specifically labeled by an affinity reagent for the acetylcholine receptor binding site. *J. Biol. Chem.* **259**, 11662–11665 (1984).
- Damle, V. N. & Karlin, A. Effects of agonists and antagonists on the reactivity of the binding site disulfide in acetylcholine receptor from *Torpedo californica*. *Biochemistry* **19**, 3924–3932 (1980).

29. Celie, P. H. *et al.* Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* **41**, 907–914 (2004).
30. Mukhtasimova, N., Free, C. & Sine, S. M. Initial coupling of binding to gating mediated by conserved residues in the muscle nicotinic receptor. *J. Gen. Physiol.* **126**, 23–39 (2005).
31. Quiram, P. A., McIntosh, J. M. & Sine, S. M. Pairwise interactions between neuronal $\alpha 7$ acetylcholine receptors and α -conotoxin Pn1B. *J. Biol. Chem.* **275**, 4889–4896 (2000).
32. Lee, W. Y. & Sine, S. M. Principal pathway coupling agonist binding to channel gating in nicotinic receptors. *Nature* **438**, 243–247 (2005).
33. Campos-Caro, A. *et al.* A single residue in the M2–M3 loop is a major determinant of coupling between binding and gating in neuronal nicotinic receptors. *Proc. Natl Acad. Sci. USA* **93**, 6118–6123 (1996).
34. Kusama, T., Wang, J. B., Spivak, C. E. & Uhl, G. R. Mutagenesis of the GABA ρ 1 receptor alters agonist affinity and channel gating. *Neuroreport* **5**, 1209–1212 (1994).
35. Lynch, J. W., Rajendra, S., Barry, P. H. & Schofield, P. R. Mutations affecting the glycine receptor agonist transduction mechanism convert the competitive antagonist, picrotoxin, into an allosteric potentiator. *J. Biol. Chem.* **270**, 13799–13806 (1995).
36. Rajendra, S. *et al.* Mutation of an arginine residue in the human glycine receptor transforms β -alanine and taurine from agonists into competitive antagonists. *Neuron* **14**, 169–175 (1995).
37. Takeuchi, A. & Takeuchi, N. A study of the action of picrotoxin on the inhibitory neuromuscular junction of the crayfish. *J. Physiol. (Lond.)* **205**, 377–391 (1969).
38. Etter, A. *et al.* Picrotoxin blockade of invertebrate glutamate-gated chloride channels: subunit dependence and evidence for binding within the pore. *J. Neurochem.* **72**, 318–326 (1999).
39. Fatima-Shad, K. & Barry, P. H. Anion permeation in GABA- and glycine-gated channels of mammalian cultured hippocampal neurons. *Proc. R. Soc. Lond. B* **253**, 69–75 (1993).
40. Sunesen, M. *et al.* Mechanism of Cl⁻ selection by a glutamate-gated chloride (GluCl) receptor revealed through mutations in the selectivity filter. *J. Biol. Chem.* **281**, 14875–14881 (2006).
41. Wada, A. The α -helix as an electric macro-dipole. *Adv. Biophys.* 1–63 (1976).
42. Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T. & MacKinnon, R. X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* **415**, 287–294 (2002).
43. Wilson, G. G., Pascual, J. M., Brooijmans, N., Murray, D. & Karlin, A. The intrinsic electrostatic potential and the intermediate ring of charge in the acetylcholine receptor channel. *J. Gen. Physiol.* **115**, 93–106 (2000).
44. Keramidas, A., Moorhouse, A. J., Schofield, P. R. & Barry, P. H. Ligand-gated ion channels: mechanisms underlying ion selectivity. *Prog. Biophys. Mol. Biol.* **86**, 161–204 (2004).
45. Mancinelli, R., Botti, A., Bruni, F., Ricci, M. A. & Soper, A. K. Hydration of sodium, potassium, and chloride ions in solution and the concept of structure maker/breaker. *J. Phys. Chem. B* **111**, 13570–13577 (2007).
46. Akabas, M. H., Kaufmann, C., Archdeacon, P. & Karlin, A. Identification of acetylcholine receptor channel-lining residues in the entire M2 segment of the α subunit. *Neuron* **13**, 919–927 (1994).
47. Lee, D. J., Keramidas, A., Moorhouse, A. J., Schofield, P. R. & Barry, P. H. The contribution of proline 250 (P-2') to pore diameter and ion selectivity in the human glycine receptor channel. *Neurosci. Lett.* **351**, 196–200 (2003).
48. Reeves, D. C., Goren, E. N., Akabas, M. H. & Lummiss, S. C. Structural and electrostatic properties of the 5-HT₃ receptor pore revealed by substituted cysteine accessibility mutagenesis. *J. Biol. Chem.* **276**, 42035–42042 (2001).
49. Barrantes, F. J. Structural basis for lipid modulation of nicotinic acetylcholine receptor function. *Brain Res. Rev.* **47**, 71–95 (2004).
50. Sherrington, C. S. *Integrative Action of the Nervous System* (Yale Univ. Press, 1906).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to H. Lester for providing the initial GluCl construct, to D. Cawley for monoclonal antibody production, to J. Michel for Fab fragment cloning and sequencing, to C. Alexander and D. C. Dawson for providing *Xenopus* oocytes, to M. Mayer for advice and equipment related to oocyte experiments, and to L. Vaskalis for help with illustrations. We thank the staff at the Advanced Photon Source beamline 24-ID-C for assistance with X-ray data collection. We are particularly appreciative of discussions with E.G. laboratory members and E. McCleskey. This work was supported by an individual NIH National Research Service Award (F32NS061404) to R.E.H. E.G. is an investigator with the Howard Hughes Medical Institute.

Author Contributions R.E.H. and E.G. contributed to all aspects of the project.

Author Information Atomic coordinates and structure factors have been deposited with the Protein Data Bank under codes 3RHW, 3RIF, 3RI5 and 3RIA for the GluCl–Fab–ivermectin complex alone and with glutamate, with picrotoxin and with iodide, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.G. (gouauxe@ohsu.edu).

METHODS

Construct design. The gene encoding the full-length *C. elegans* GluCl α protein (Genbank accession code AAA50785.1)⁸, including the native signal peptide and a C-terminal 8 \times -histidine tag, was codon optimized and subcloned into the pFastBac1 vector for baculovirus-driven expression in Sf9 insect cells. A construct for FSEC-based⁹ small-scale screening of detergent stability, mutagenesis and purification additionally contained the enhanced GFP (EGFP)-coding sequence inserted into the M3–M4 loop region as previously described^{45–53}. To improve crystallization behaviour, 41 amino acid residues from the N terminus and 6 from the C terminus were removed, and residues K345–K402 (in the mature, full-length sequence), corresponding to the M3–M4 loop, were substituted with the residues AGT.

GluCl expression and purification. Bacmid and baculovirus were generated from pFastBac1 constructs and Sf9 cells were infected at 27 °C using standard methods. After 18 h of infection, cells were maintained shaking at 20 °C, and then harvested for purification after 72–96 h. Cells were collected by centrifugation at 6,200g and disrupted using an EmulsiFlex-C5 (Avestin) in buffer containing 20 mM Tris pH 7.4, 150 mM NaCl (TBS buffer), and 1 mM PMSF. The homogenate was clarified by centrifugation at 9,700g, and crude membranes were collected from the light membrane fraction by centrifugation at 125,000g. The membranes were mechanically homogenized and solubilized in 0.25 g C₁₂M (*n*-dodecyl-B-D-maltopyranoside; Anatrace) per gram of membranes in TBS. Solubilized membranes were centrifuged at 125,000g. Supernatant containing GluCl_{cryst} was bound to TALON Co²⁺-affinity resin (Clontech), washed with TBS solution containing 1 mM C₁₂M and 25 mM imidazole, and eluted with 250 mM imidazole. All purification steps were performed at 4 °C.

Monoclonal antibody generation and Fab purification. The mouse monoclonal antibody against GluCl (IgG1, λ) was obtained using standard methods⁵⁴. Specificity of the antibody for properly folded pentameric GluCl_{cryst} was assayed by FSEC and western blot. Cloning and sequencing of Fab antibody regions were performed from mouse hybridoma cells. Antibody was purified from hybridoma supernatants by cation exchange and protein A affinity chromatography. Fab fragments were generated by papain digest of whole antibody, and purified by protein A chromatography to remove Fc molecules and undigested material, followed by anion exchange.

Purification of GluCl_{cryst}–Fab complex. Eluent from Co²⁺-affinity purification and Fab from ion exchange were mixed to an excess of Fab to GluCl_{cryst} subunits, concentrated, and applied to a gel filtration column (Superose 6 10/300 GL, GE Healthcare Life Sciences) equilibrated in TBS + 1 mM C₁₂M. GluCl_{cryst}–Fab complex was concentrated to 1–2 mg ml^{−1}. For samples used in crystallization, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) or 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) lipids (Avanti Polar Lipids) were added to 0.02% from a 2% stock suspension in 20% DMSO, 80% gel filtration buffer, and ivermectin (Sigma) was added to 0.1 mM from a 10 mM stock in DMSO.

Crystallization and cryoprotection. Initial crystallization attempts of GluCl constructs in the absence of Fab resulted in poorly diffracting (8 Å) crystals that grew in a very limited range of crystallization conditions. Crystallization of the Fab complex occurred in diverse conditions; best diffracting crystals were obtained in hanging-drop format and diffracted to 4–5 Å. Crystals diffracting beyond 4 Å were obtained only in the presence of Fab, either POPC or DPPC, and ivermectin. These tetragonal crystals grew by vapour diffusion at 4 °C in 21–23% PEG 400, 50 mM sodium citrate pH 4.5, and 70 mM sodium chloride, and diffracted maximally to Bragg spacings of 3.26 Å (Supplementary Table 1). Crystals were protected before flash freezing in liquid nitrogen by 1–2 min soaks in crystallization solution supplemented to contain 30% PEG 400. To obtain structures of GluCl_{cryst} in complex with additional ligands, crystals of the same form were soaked briefly in cryoprotectant containing either 5 mM picrotoxin (picrotoxinin, the more active component of picrotoxin, was used, obtained from Sigma), 50 mM L-glutamate or 300 mM sodium iodide. In an effort to minimize occupancy of chloride in the iodide-soaked crystals, crystals were transferred serially into three replicate cryoprotectant solutions lacking chloride before flash freezing. Nonetheless, because the iodide soaks were only 1–2 min, some chloride may have been carried over from crystallization. Electron density maps derived from these crystals showed no significant change in the strength of the electron density feature in the pore where we have modelled a chloride ion. We also soaked crystals in an analogous manner in bromide-containing cryo-solutions but were not able to observe significant peaks in the resultant anomalous difference electron density maps.

Data collection. Diffraction data were collected using synchrotron radiation at the Advanced Photon Source (Argonne National Laboratory, beamline 24-ID-C) with a mini-Kappa goniometer and in-house crystal alignment strategy software. The best-ordered crystals have a diffraction limit of 3.26 Å, a mosaic spread of 0.2–0.5°, and they are of the space group P4₃2₁2 with one GluCl_{cryst}–Fab complex

per asymmetric unit. The unit cell dimensions are $a = b = 155$ Å, $c = 575$ Å, $\alpha = \beta = \gamma = 90^\circ$, resulting in a Matthews' coefficient (V_M) of 4.0 Å³ Da^{−1} (ref. 55). Diffraction data were indexed, integrated and scaled using HKL2000 (ref. 56) or Xia2 (refs 57–62) software.

Structure determination. The structure was solved by molecular replacement using Phaser⁶³; the search probe was a pentameric homology model of GluCl_{cryst} made from GLIC (PDB code 3EHZ)¹⁰, using Swiss-Model⁶⁴. After an initial solution was found, phases were improved by solvent flattening⁶⁵ and electron density for Fab molecules bound at each of the five subunit interfaces of GluCl_{cryst} became plainly visible. A Fab homology model was made, using PDB 1NGQ for the light chain and 1F3D for the heavy chain, and Coot⁶⁶ to overlay the two modelled chains to make a single Fab molecule. Fab CDR loops were truncated and the model was used for molecular replacement using the GluCl_{cryst} solution as a starting point. In this manner, a single Fab was placed, and by copying the remaining Fab molecules around the fivefold non-crystallographic symmetry (NCS) axis, approximate positioning of all Fab molecules was accomplished. Electron density for Fab constant domain regions was poor after NCS averaging, and from non-averaged maps it was clear that the Fab constant domains did not obey five-fold symmetry. A starting model that included GluCl_{cryst} and five Fab variable domains was used for automated building with Buccaneer⁶⁷. Electron density maps were then good enough to position ivermectin molecules in the transmembrane domain loci, and to begin manual building of the Fab constant domains. Ivermectin stereochemistry, determined previously, is modelled as such⁶⁸. Numbering of ivermectin atoms in the figures is as defined in the PDB files; Supplementary Table 2 relates this numbering to that from the small molecule structure.

Iterative refinement of the model against the X-ray data using Phenix⁶⁹, manual adjustment in Coot into simulated annealing composite omit electron density maps⁶⁹ or real-space averaged maps⁷⁰, and structure quality analysis using Molprobity⁷¹ were carried out until satisfactory model statistics were obtained. Three groups of fivefold NCS restraints were present during refinement: five subunits of GluCl_{cryst}, five heavy chain Fab variable domains (residues 1–120), and the five light chain Fab variable domains (residues 1–108); the root mean squared deviation (r.m.s.d.) values between the chains within each of these three groups were 0.017, 0.014 and 0.015 Å, respectively. Isotropic B factors and TLS parameters were also refined; the 15 TLS groups comprised five GluCl_{cryst} subunits, five Fab variable domains, and five Fab constant domains. The final models contain the GluCl_{cryst} pentamer from residues 1–339 or 340, five ivermectin molecules, a single N-linked carbohydrate at N185 in three of the five subunits, five Fab molecules (1–221 for heavy chains, 1–210 for light chains), and several lipid and detergent molecules. Some portions of the Fab constant domains lacked electron density in composite omit maps and hence were omitted from the final model. The iodide-bound structure is of very low resolution and not completely refined: several anomalous difference electron density peaks in the extracellular domain were not modelled with iodide atoms.

Sequence alignments were made using PROMALS3D (ref. 72) and ClustalW⁷³. Isoelectric surface calculations were made using the APBS⁷⁴ add-on in PyMOL⁷⁵. Pore dimensions were analysed using HOLE software⁷⁶.

Electrophysiology. RNAs encoding GluCl proteins were transcribed from pGEM-HE⁷⁷ plasmids using the mMessage mMachine T7 Ultra kit (Ambion). Defolliculated stage V–VI *Xenopus* oocytes were provided by D. C. Dawson and C. Alexander, prepared as previously described⁷⁸. Oocytes were injected with 25 ng of GluCl RNAs, and current recordings were made 3–5 days afterwards. Frog saline (FS) recording solution contained (in mM): 96 NaCl, 2 KCl, 1 MgCl₂, 1.8 CaCl₂, 5 HEPES pH 7.5. Recording solution for iodide permeability experiments was FS but with NaI in place of NaCl. All ligands were made up in FS from stock solutions in water, except: picrotoxin, 1 M stock in DMSO; ivermectin, 5 mM stock in DMSO. Recording electrode pipettes (0.7–2 M Ω) were cushioned with 0.8% LMP agarose in 3 M KCl and backfilled with 3 M KCl. Oocytes were voltage-clamped at −80 mV except in experiments to determine the reversal potential, which used 40 ms voltage steps. Analogue data were filtered at 50 Hz and digitized at ≥ 1 kHz. The Axoclamp 2B amplifier (Axon Instruments) and pClamp 10 software (Molecular Devices) were used for data acquisition. In un-injected oocytes, no significant responses to test solutions were observed (Supplementary Fig. 13).

Radioligand binding experiments. Experiments to test binding of [³H]-L-glutamate to GluCl_{cryst} and competition of the radioligand with other compounds were done using purified Nano15-tagged⁷⁹ GluCl_{cryst} (N-terminal tag) and streptavidin-Ysi scintillation proximity assay beads (SPA; GE Healthcare Life Sciences). The concentration of binding sites was fixed at 100 nM after a preliminary experiment to determine optimal GluCl_{cryst} concentration (Supplementary Fig. 12). Other binding assay components were: 50 mM Tris pH 7.4, 150 mM NaCl, 1 mM C₁₂M, 1 mg ml^{−1} SPA beads, and 1 μ M ivermectin. Saturation

binding of [^3H]-L-glutamate in the presence and absence of Fab was performed with a 1:30 dilution of specific activity of the radiolabel with [^1H]-L-glutamate, and a slight molar excess of Fab to GluCl_{cryst} subunits as verified by FSEC experiments. Measurement of background signal in saturation binding experiments was complicated by, we believe, significant binding of [^3H]-L-glutamate directly to SPA beads and lack of a chemically distinct competitor for the neurotransmitter-binding site. Neither high concentrations of [^1H]-L-glutamate or absence of protein were able to fully account for this apparently non-specific signal. To address the background component that was not accurately measured experimentally, we combined subtraction of a background signal measured in the absence of GluCl_{cryst} with a linear component still present in the binding data (calculated using the total binding function in the fitting software). In saturation binding experiments in the presence of Fab (Supplementary Fig. 12), data were better fit after removing background signal measured in the presence of 10 mM [^1H]-L-glutamate combined with the calculated linear component. In competition binding experiments to determine IC₅₀ values, [^3H]-L-glutamate was 1 μM using a 1:10 dilution of specific activity of the radiolabel with cold glutamate. In all [^3H]-L-glutamate and electrophysiological dose–response experiments, data were fit with GraphPad Prism software.

51. Gensler, S. *et al.* Assembly and clustering of acetylcholine receptors containing GFP-tagged ϵ or γ subunits: selective targeting to the neuromuscular junction *in vivo*. *Eur. J. Biochem.* **268**, 2209–2217 (2001).
52. Li, P., Slimko, E. M. & Lester, H. A. Selective elimination of glutamate activation and introduction of fluorescent proteins into a *Caenorhabditis elegans* chloride channel. *FEBS Lett.* **528**, 77–82 (2002).
53. Slimko, E. M. & Lester, H. A. Codon optimization of *Caenorhabditis elegans* GluCl ion channel genes for mammalian cells dramatically improves expression levels. *J. Neurosci. Methods* **124**, 75–81 (2003).
54. Harlow, E. & Lane, D. *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1988).
55. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497 (1968).
56. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
57. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
58. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82 (2006).
59. Leslie, A. G. W. Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography, No. 26 (1992).
60. Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **62**, 48–57 (2006).
61. Sauter, N. K., Grosse-Kunstleve, R. W. & Adams, P. D. Robust indexing for automatic data collection. *J. Appl. Cryst.* **37**, 399–409 (2004).
62. Zhang, C. Y., Sauter, N. K., van den Bedem, H., Snell, G. & Deacon, A. M. Automated diffraction image analysis and spot searching for high-throughput crystal screening. *J. Appl. Cryst.* **39**, 112–119 (2006).
63. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
64. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201 (2006).
65. Cowtan, K. Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography **31**, 34–38 (1994).
66. Emsley, P. & Cowtan, P. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
67. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62**, 1002–1011 (2006).
68. Springer, J. P., Arison, B. H., Hirshfield, J. M. & Hoogsteen, K. The absolute stereochemistry and conformation of avermectin B_{2a} aglycon and avermectin B_{1a}. *J. Am. Chem. Soc.* **103**, 4221–4224 (1981).
69. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
70. Collaborative Computational Project 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–776 (1994).
71. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
72. Pei, J., Kim, B. H. & Grishin, N. V. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295–2300 (2008).
73. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
74. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**, 10037–10041 (2001).
75. DeLano, W. L. *The PyMOL Molecular Graphics System* (<http://www.pymol.org>) (DeLano Scientific, 2002).
76. Smart, O. S., Goodfellow, J. M. & Wallace, B. A. The pore dimensions of gramicidin A. *Biophys. J.* **65**, 2455–2460 (1993).
77. Liman, E. R., Tytgat, J. & Hess, P. Subunit stoichiometry of a mammalian K⁺ channel determined by construction of multimeric cDNAs. *Neuron* **9**, 861–871 (1992).
78. Alexander, C. *et al.* Cystic fibrosis transmembrane conductance regulator: using differential reactivity toward channel-permeant and channel-impermeant thiol-reactive probes to test a molecular model for the pore. *Biochemistry* **48**, 10078–10088 (2009).
79. Lamla, T. & Erdmann, V. A. The Nano-tag, a streptavidin-binding peptide for the purification and detection of recombinant proteins. *Protein Expr. Purif.* **33**, 39–47 (2004).